Epistemic Risk and the Demands of Rationality

Richard Pettigrew

March 8, 2021

Abstract

The short abstract

Epistemic utility theory +
permissivism about attitudes to epistemic risk ⇒
permissivism about rational credences.

The longer abstract I argue that epistemic rationality is permissive. More specifically, I argue for two claims. First, a radical version of interpersonal permissivism about rational credence: for many bodies of evidence, there is a wide range of credal states for which there is some individual who might rationally adopt that state in response to that evidence. Second, a slightly less radical version of intrapersonal permissivism about rational credence: for many bodies of evidence and for many individuals, there is a narrower but still wide range of credal states that the individual might rationally adopt in response to that evidence.

My argument proceeds from two premises: (1) epistemic utility theory; and (2) permissivism about attitudes to epistemic risk. Epistemic utility theory says this: What it is epistemically rational for you to believe is what it would be rational for you to choose if you got to pick your beliefs and, when picking them, you cared only for their purely epistemic value. So, to say which credences it is epistemically rational for you to have, we must say how you should measure purely epistemic value, and which decision rule it is appropriate for you to use when you face the hypothetical choice between the possible credences you might adopt. Permissivism about attitudes to epistemic risk says that rationality permits many different attitudes to epistemic risk. These attitudes can show up in epistemic utility theory in two ways: in the way that you measure epistemic value; and in the decision rule that you use to pick your credences. I explore what happens if we encode our attitudes to epistemic risk in our epistemic decision rule. The result is the interpersonal and intrapersonal permissivism described above: different attitudes to epistemic risk lead to different choices of priors; given most bodies of evidence you might acquire, different priors lead to different posteriors; and even once we fix your attitudes to epistemic risk, if they are at all risk-inclined, there is a range of different priors and therefore different posteriors they permit. The essay ends by considering a range of objections to the sort of permissivism for which I've argued.

Contents

1 Introduction

It is good to be rational. It is better to think and reason and believe and act rationally than to do so irrationally. The rationality that governs what we believe and how strongly, and the reasoning that leads us to those beliefs, we call *theoretical* or *epistemic rationality*. This distinguishes it from *practical* or *pragmatic* or *prudential rationality*, which governs how we act. Epistemic rationality will be my topic here. And, in particular, I want to know what epistemic rationality demands of us. More specifically still, I want to know: how strict are its demands? We can, of course, agree that it will make different demands of people who have different evidence. It might demand that I believe I had Countess Grey tea with breakfast this morning, because my evidence includes memories of doing exactly that; but, until I told you, it might not have demanded the same of you, since perhaps you didn't have any evidence one way or the other. But once we fix the evidence we have, what then does it demand? Is there always a single unique rational response to that evidence? Or are there bodies of evidence I might have to which there are a number of different rational responses? That's the question of this essay. My answer is that there is. Epistemic rationality, I will argue, is permissive. In fact, I'll argue that it is radically so.

As I write this, we are two days out from the US Presidential election of 3rd November 2020 to choose between the incumbent, Donald Trump, the Democratic Party's nominee, Joe Biden, and the Libertarian Party's nominee, Jo Jorgensen. A handful of states have yet to be called: Arizona, Nevada, Georgia, and Pennsylvania. Biden could win with Pennsylvania alone, or with Arizona and Nevada together, or with some other less likely combination. To a decent approximation, I have the same evidence as most of my friends. We are all watching the same data flow in from the counts through the same news websites. But our reactions are markedly different. Some believe Biden will win, while others withhold judgment. Some have very high confidence that Biden will win—close to certainty, in fact. Some assign a more middling credence to that eventuality. Of course, once the result is finally determined, some will be right in their beliefs and some will be wrong, and some will have more accurate credences, and some will have less. But, at the moment, you might think that rationality permits each response.

One way to argue for this is to note that the evidence that I share with my friends is complex. Different parts of it point in different directions. Professional election-watchers have told us the likely composition of the districts whose votes remain to be counted, and that points towards a Biden victory. But the past 48 hours have seen their other predictions refuted again and again, and that makes us trust them less. Alongside that detailed, local, particular information, I have some more general information:

for instance, I know that it's very unusual for the Democrats to win back the Presidency after the Republicans have held it for only a single term—it hasn't happened since Grover Cleveland won against Benjamin Harrison in the election of 1892. And alongside that, I also know that Trump will try hard to challenge any count that favours Biden—he's all but said he will. How should I combine these different parts of my evidence to determine my beliefs? You might well think that there are different ways to weigh and combine these disparate parts of my evidence to determine my attitudes to the three possibilities. And you might think that each is rationally permissible. And, if you do, you'll probably think that the different attitudes that result from these different ways of combining the evidence are all permitted by rationality.[1]

This is one motivation for a permissive conception of epistemic rationality. When evidence is complex, it says, we must weigh parts of it against each other to determine our beliefs. And, the argument goes, there is no privileged way to do that. Therefore, rationality permits any number of ways of doing it, and so it permits the many different outcomes of those different ways of weighing evidence. What's more, even if there were a unique best way to weigh the different parts of our evidence and combine them, it might be that it is cognitively very demanding to follow it to the letter. And in that case, you might think that rationality doesn't require you to follow it perfectly, but only to approximate it to some degree. If imperfectly following the unique best way to combine your evidence were then to result in different attitudes, those different attitudes might all be rationally permissible.[2]

So rationality might be permissive about how we should respond to complex evidence. It might also be permissive about how we should respond to very sparse evidence. It is my six year old godson's first time watching a US presidential election. He knows that there are three people in the running and he knows their names; but he knows almost nothing more. Sometimes, he's a little more confident that Joe will win, sometimes a little more confident that Donald will; sometimes, he plumps for Jo. Whereas some parts of my evidence point in one direction and other parts point in another, his evidence doesn't point in any direction. Again, it seems that rationality is permissive. Each of his responses seems rational.[3]

[1] Schoenfield (2019) and Titelbaum & Kopec (ms, Section 4) offer arguments for permissivism of this sort.

[2] Jackson (2019) develops an argument of this sort.

[3] Elga (2010) offers another example of sparse evidence that seems to determine no unique rational response:

> A stranger approaches you on the street and starts pulling out objects from a bag. The first three objects he pulls out are a regular-sized tube of toothpaste, a live jellyfish, and a travel-sized tube of toothpaste. To what degree

Of course, we might say that, absent any evidence that tells them apart, he should be equally confident in each of them winning. But that depends on taking the three-way partition of the possibilities to be the relevant one. After all, he might first divide up the possibilities into those in which Joe wins and those in which Joe loses, and divide his credences equally over those; and only then take the credence he assigns to Joe losing and divide it equally between Donald winning and Jo winning. Does rationality really demand that he divide his credences equally over the options in one partition rather than the options in another? Remember: he knows nothing more than that the three of them are running.[4]

This is another motivation for a permissive conception of epistemic rationality. When evidence is sparse, it says, there are many ways to respond to it that respect that evidence, and there is no privileged way to choose between them. Therefore, rationality permits any number of ways of choosing, and so it permits the many different outcomes of those different ways.

In this essay, I'd like to offer a different reason for thinking that epistemic rationality is permissive. The argument begins with a thesis about what determines the rationality of doxastic attitudes, such as beliefs and credences. The thesis is sometimes called *epistemic utility theory*.[5] It is what Selim Berker (2013) calls a teleological theory in epistemology. That is, according to epistemic utility theory, the epistemic right—that is, what it is rational to believe—depends on the epistemic good—that is, what it is valuable to believe, epistemically speaking. Thus, according to epistemic utility theory, beliefs and credences can have more or less purely epistemic value given different ways the world is. For instance, according to the version of the theory known as *accuracy-first epistemology*, which combines epistemic utility theory with the axiological thesis of *veritism*, a belief or credence is more valuable, epistemically speaking, the more accurately it represents the world. Thus, a belief or higher credence in a truth is more valuable than a disbelief or lower credence in it; and a disbelief or lower credence in a falsehood is more valuable than a belief or higher credence in it.

To the claim that doxastic attitudes, like beliefs and credences, have purely epistemic value, epistemic utility theory adds the claim that which attitudes are rationally permissible is determined by facts about this epis-

> should you believe that the next object he pulls out will be another tube of toothpaste? (Elga, 2010, 1)

[4]In (Pettigrew, 2016b), I offer an argument for impermissivism in this case: I argue that my godson should divide his credences equally over the most fine-grained partition that his conceptual scheme affords him. As we'll see below, I reject that argument now.

[5]Some representative pieces that deploy the version that I will explore: (Oddie, 1997; Joyce, 1998; Greaves & Wallace, 2006; Joyce, 2009; D'Agostino & Sinigaglia, 2010; Williams, 2012; Easwaran, 2013; Schoenfield, 2015, 2016; Pettigrew, 2016a; De Bona & Staffel, 2017; Schoenfield, 2017).

temic value, given different ways the world might be, together with bridge principles from decision theory that allow us to infer facts about rationality from facts about value. Decision theory tells us what we are rationally permitted to do when we face a given decision problems. A *decision problem* consists of a range of available options. In the practical case, these are different actions we might perform, such as voting for a particular presidential candidate rather than another, consenting to one medical treatment from a range of alternatives, putting on a scarf or a hat or a winter coat when we leave the house, and so on. And a *decision rule* takes such a decision problem, together with some other ingredients, and determines which of the available options are rational for you to choose when faced with that problem. The other ingredients nearly always include the utilities that you assign to the options, given different ways the world might be. These are numerical measures of the values that you assign to those options at those different possible worlds. But they might also include: objective probabilities over the different possible worlds; subjective probabilities over those worlds, or credences, as we're calling them here; attitudes to risk; and perhaps other items. It is these decision rules that supply the bridge principles that allow us to derive facts about epistemic rationality from facts about epistemic value. According to epistemic utility theory, the rationally permissible doxastic attitudes for an individual are determined by: (i) considering the decision problem in which the available options are not the actions the individual might perform, but rather the different possible doxastic attitudes they might adopt, and the utilities are not pragmatic utilities, but epistemic utilities; and (ii) applying the appropriate decision rule to this decision problem. So the epistemic utilities encode the goal of having credences, while the decision rule specifies the rationally permissible ways to pursue that goal. The rational credences are the rationally permissible ways to pursue the goal of epistemic value.

So that is epistemic utility theory, and it provides the framework in which our argument takes place. To this, I add the claim, inspired by William James' 1897 essay 'The Will to Believe', that many different attitudes to epistemic risk are permissible (James, 1897). James, thinking mainly of the case of belief, thought that there were two competing epistemic goals: Believe truth! Shun error! The epistemically risk-inclined place greater weight on the former; the epistemically risk-averse weight the latter more heavily. Both, James thinks, are permissible—which you have is a matter of subjective preference.

Following my own earlier treatment of James' epistemology in epistemic utility theory, I note that there are two ways to make James' claim precise (Pettigrew, 2016c). In applications of decision theory to practical choices, there are two ways to encode attitudes to risk: you can encode them in your utility function or in your decision rule. In Section 3.1, I describe Thomas Kelly's Jamesian argument for permissivism about rational belief, which proceeds by encoding attitudes to epistemic risk in the epistemic utility function and showing that different permissible attitudes to epistemic risk, so encoded, give rise to different permissible beliefs given the same evidence (Kelly, 2014). In Section 3.2, I note Sophie Horowitz's observation that we cannot hope for a similar argument in the credal case (Horowitz, 2017).

In Chapter 4, I argue that we should, instead, encode our attitudes to epistemic risk in the decision rule we apply to certain epistemic decision problems. In particular, I argue that the decision rule we use to pick our prior credences—those we have before we acquire any evidence—is different from the rule we use to pick our posterior credences—those we have after evidence has arrived. To pick our priors, we must use a decision rule that does not include any probabilities among the ingredients needed to apply it—after all, at the point when we're picking our priors, there are no probabilities to hand. I characterise such a family of such rules in Chapter 4: it is the *generalised Hurwicz Criterion* (GHC), and it includes rules that encode a wide variety of attitudes to risk. As we see in Chapter 5, if, like James, we are permissive about rational attitudes to epistemic risk, this family of decision rules delivers radical interpersonal and moderate intrapersonal permissivism for the prior credences chosen using it. Then, in Chapter 6, I argue that you should pick your posteriors using your prior credences and a risk-neutral decision rule. And, building on work that Hannes Leitgeb and I undertook over a decade ago, as well as a more recent development by Dmitri Gallow, I argue that you should obtain your posteriors from your priors via Bayes' Rule, or Bayesian Conditionalization (Leitgeb & Pettigrew, 2010b; Gallow, 2019). Combined with the radical interpersonal and moderate intrapersonal permissivism for priors for which we argued before, this gives radical interpersonal and moderate intrapersonal permissivism for posteriors as well.

I conclude, in Chapter 7, by considering how the argument just described helps us answer five objections to permissivism.

A number of themes recur throughout the essay. First, we will see at a number of points that Jamesian permissivism about epistemic risk, and the permissivism about rational credences that it entails, is a rather fragile creature. Quite often, I will adopt principles, norms, and theses that, if strengthened only slightly, either entail impermissivism or plunge us into inconsistency. Of course, I will explain why we should go so far as to adopt the principles in question, but not further. But it is nonetheless of interest that our position might collapse in a number of places either into the impermissivism we seek to avoid, or into outright inconsistency.

Second, as will become most obvious in the responses I offer to some of the objections in Chapter 7, when we debate permissivism about epistemic rationality, it is crucial that we say what we might mean by rationality in general and in the epistemic case in particular. Epistemic utility theory offers a very specific account of epistemic rationality: the epistemic utility functions provide the legitimate ways of measuring what you value, epistemically, about your doxastic attitudes; and the decision rules provide constraints on the rational ways to pursue the goal of epistemic value, so measured. As we will see when we respond to some of the general objections to permissivism below, when we have this specific account in hand, we can see why some of the general principles on which these objections rely simply do not hold.

Finally, we will use formal tools throughout. I find myself torn about this. On the one hand, I think they provide an important methodology in contemporary philosophy. We pride ourselves on making our claims precise and identifying imprecision in the claims of others. Formal tools cannot do this on their own, and indeed they can often obscure meaning and intention; but, used judiciously, they can also assist greatly as we strive for greater clarity and transparency. In the present essay, they allow me to specify my position precisely enough that we can determine its consequences, reassure ourselves that it is internally consistent, and tease out nuances that would otherwise remain hidden. As I mentioned above, what follows will reveal that it is not straightforward to provide a consistent version of a Jamesian epistemology that does not collapse into impermissivism. That fact itself only becomes apparent when we formalise our approach. Nonetheless, I am aware that these methods can deter some readers. I've tried to make much of the technical material self-contained, though I will at various points refer elsewhere for major results. But I know that's not sufficient. We're all busy people with finite time, and it would be arrogant of me to think that my ideas are so valuable that it is worth the time to learn all of the technical machinery from scratch in order to get hold of the philosophical ideas. So, at the end of each chapter, I will provide an overview of the argument of that chapter that does not use formal techniques (the blue boxes headed 'In brief...'). As well as providing a route through the argument for those less comfortable with the formal machinery, these summaries provide a way to discern the overall shape of the argument reasonably quickly without delving into its details.

2 Varieties of Permissivism

Epistemic rationality is permissive. That's our claim. But, as often with philosophical claims, when you start making it precise, it fractures into a number of more fully specified theses. The rough taxonomy that follows is inspired by Elizabeth Jackson's (2019) and extends it a little.

2.1 What sort of doxastic state?

First, there are different versions of the claim for different sorts of doxastic attitudes we might have in response to evidence. Let's run through some of those different sorts of attitude.

We might have *beliefs*. These are sometimes called *full beliefs*, *all-or-nothing beliefs*, or *on-off beliefs*. They are categorical states. They are what you report when you say, 'I hold that anthropogenic climate change is real', or 'I think the butter is in the fridge', or 'I believe that God exists'. We represent your total belief state by the three sets: the set of propositions you believe, the set you disbelieve, and the set on which you withhold judgment.

We might have *precise credences*. These are sometimes called *partial beliefs*, *degrees of belief*, or *subjective probabilities*. They are graded states. They are what you report when you say, 'I'm 80% sure it's the next turning on the right' or 'I think it's 50% likely that Serena will win' or 'I'm only 5% confident that I will see my parents this year'. We represent your credal state by a function that takes each proposition about which you have an opinion and assigns to it your precise credence in that proposition. That credence is represented by a real number that is at least 0 and at most 1. The stronger the belief, the higher the number that represents it. It measures how strongly you believe the proposition. 0 represents minimal credence, while 1 represents maximal credence. That function is known as your *credence function*.

We might have *comparative confidences*.[6] These also represent the strength of our belief in different propositions, but they don't do it numerically. Rather, they order the propositions about which we have an opinion by how strongly we believe them, allowing that some propositions might not be comparable. We report them when we say, 'It's more likely that Sarah and Telv will stay married than John and Melissa'.

We might have *imprecise credences*. These are sometimes called *mushy credences* or *imprecise probabilities*.[7] These represent our credal state not by a single precise numerical credence function, but by a set of them, which we call our *representor* (van Fraassen, 1990). A natural interpretation is this:

[6] Since these are less familiar doxastic states, let me recommend an overview for the interested reader: (Konek, 2019).

[7] Again, let me recommend this overview: (Bradley, 2016).

what is determinately true of our doxastic state is what is true according to every credence function in the representor. So, if I think that rain in Bristol tomorrow is more likely than sun in Bristol tomorrow, every credence function in my representor assigns higher credence to the former than the latter. But if I don't think rain more likely than sun, nor sun more likely than rain, nor both equally likely, then my representor contains some credence functions assign higher credence to rain, some assign higher credence to sun, and some assign the same credence to them both. Imprecise credences capture well what I report when I say, 'I'm between 50% and 80% more confident that it will rain by this afternoon'.

And so on. There are other ways to represent doxastic states, such as *ranking functions* (Huber, 2006). But the examples we have given will be sufficient.

For each way of representing a doxastic state, there is a different version of permissivism about epistemic rationality. For instance, according to the precise credal version of permissivism, there is evidence we might obtain that doesn't determine a unique rational credence function on the set of propositions about which you have an opinion. According to the imprecise credal version, there is evidence that doesn't determine a unique representor. And so on.

For many different sorts of doxastic attitude, you can be permissive with respect to one but not the other. For instance, take my complex evidence concerning the outcome of the 2020 US Presidential election. You might think that there is no precise credence in a Biden win that rationality requires me to adopt. After all, what could determine that 0.787 is permissible, for instance, but not 0.786? But that doesn't prevent there being a unique imprecise credence that is rationally required. It might, for instance, be the set of all the rationally permissible precise credences.

However, you might find it harder to see how rationality could be permissive about imprecise credences but not about precise ones. Indeed, since precise credences are usually seen as a particular case of imprecise credences—a precise credence function represents the same doxastic state as the representor that contains only that credence function—if rationality requires a unique precise credence function, surely it also requires the corresponding representor?

But perhaps not. Rationality might sometimes demand a unique precise credence function because of the more limited representational possibilities they afford, while permitting a number of different imprecise credal states. For instance, suppose I know that the urn in front of me contains ten balls, and they are all either green or purple. But I don't know how many of each; indeed, I have no evidence that bears on that question. Then there are eleven possibilities: zero green balls and ten purple; one green and nine purple; and so on. Now consider the proposition that the next ball drawn from the urn will be green, as well as its negation, which says that the next ball drawn will be purple. We might imagine that rationality requires that I treat the green and purple possibilities symmetrically, but nothing beyond this. Then, among precise credence functions, rationality would permit the one that assigns $\frac{1}{2}$ to green and $\frac{1}{2}$ to purple. And it would permit only that one. Among imprecise credal states, rationality would permit the representor that includes only that precise credence function that assigns $\frac{1}{2}$ to green and $\frac{1}{2}$ to purple. But it would also permit the representor that includes the eleven precise credence functions that assign credence $0, \frac{1}{10}, \ldots, \frac{9}{10}, 1$ to green and $1, \frac{9}{10}, \ldots, \frac{1}{10}, 0$ to purple, respectively. And it would permit the representor that includes the credence functions that assign $0, \frac{1}{2}, 1$ to green and $1, \frac{1}{2}, 0$ to purple, respectively. And so on. And if that is the case, then rationality is impermissive with respect to precise credences but permissive with respect to imprecise ones.

There is lively debate about how our credences should relate to our beliefs (Kyburg, 1961; Foley, 1992; Arló-Costa & Pedersen, 2012; Buchak, 2014; Leitgeb, 2014; Staffel, 2016). One straightforward view is the so-called *Lockean thesis*. This says that there is some threshold $\frac{1}{2} \leq t < 1$ such that you should believe a proposition if your credence in it lies above t, you should disbelieve it if your credence lies below $1 - t$, and you should withhold judgment on it otherwise. Now, you might think that rationality requires each individual to set their own Lockean threshold, but it doesn't require them to pick any particular one. If this is the case, rationality might demand a unique precise credence function, without demanding unique sets of beliefs and disbeliefs and suspensions, since it does not specify a unique Lockean threshold. Similarly, even if it specifies a unique Lockean threshold, it might specify unique sets of beliefs and disbeliefs and suspensions, but not a unique credence function, since typically there will be many precise credence functions that all give rise to those beliefs and disbeliefs and suspensions via the Lockean thesis.

In this essay, I'm interested primarily in permissivism about the epistemic rationality of credences. I'll argue that there is evidence we might have for which rationality does not specify a unique rational credal response.

2.2 Interpersonal or intrapersonal?

So we're interested in credal permissivism. But this comes in two varieties: an interpersonal version and an intrapersonal version. According to interpersonal permissivism, it is possible for different rational individuals to have the same evidence, but different attitudes in response. According to the intrapersonal version, there is evidence that a single individual might have, and different doxastic attitudes such that, whichever of these attitudes they have in response to the evidence, they'll be rational.

If you think rationality is interpersonally impermissive, it's hard to see how you could think it is intrapersonally permissive. After all, if there are two attitudes and it's rationally permissible for you to adopt either in response to your evidence, it's hard to see how it could not also be rational for someone else with your evidence to adopt either, regardless of which you adopt. But there might be pathological cases. I'll leave that possibility aside.

It is, however, reasonably easy to see how rationality might be intrapersonally impermissive, but interpersonally permissive. For instance, it might be that, while a body of evidence on its own does not fix the rational response, that body of evidence combined with some feature of an individual does. Let's meet two such features that have been proposed: (i) the individual's conceptual scheme, and (ii) their inductive methods.

Recall my godson. He knows that the Presidential race is between Joe, Donald, and Jo. Suppose he divides his credence equally between Joe winning and Joe losing, and then he divides his credence in Joe winning equally between Donald winning and Jo winning. We might say that he is irrational because rationality requires him to divide his credence equally over the finest-grained set of possibilities that his conceptual scheme affords him, and that requires him to divide his credences equally over Joe winning, Jo winning, and Donald winning (Pettigrew, 2016b). On this account of rationality's requirements, we say, then, that they depend both on your evidence—in my godson's case, nothing beyond the fact that one of the three must win—and your conceptual scheme—in my godson's case, a conceptual scheme rich enough to represent all three possibilities. But you might think that rationality does not constrain which conceptual scheme you have. For my godson, then, rationality requires a credence of $\frac{1}{3}$ that Joe will win, while for someone with the same evidence, but a conceptual scheme that only affords them two possibilities—Joe winning and Joe losing—rationality requires a credence of $\frac{1}{2}$ that Joe will win. In this case, credal rationality is intrapersonally impermissive, because your evidence and your conceptual scheme determine the credence you should have; but it is interpersonally permissive, because it doesn't determine what conceptual scheme you should have.

Or, as I mentioned above, we might think that what your evidence doesn't determine a unique rational response on its own, but it does when coupled with the inductive methods that you endorse (Schoenfield, 2019; Titelbaum & Kopec, ms). According to this view, there are many different inductive methods that rationality permits you to endorse, and indeed many that will issue in different conclusions based on the same evidence—hence, interpersonal permissivism. But, fed any body of evidence, a particular inductive method will issue in a single mandated doxastic response—hence, intrapersonal impermissivism.

2.3 Radical or not?

To fix a version of permissivism, we must fix the attitudes to which it applies and whether it applies interpersonally or intrapersonally. But we must also say how radical it is. That is, we must say how varied are the different attitudes that rationality permits us to adopt in response to a given body of evidence. For instance, on a radical version of permissivism about belief, there are bodies of evidence to which rationality permits you to respond by believing a particular proposition, suspending judgment on it, or disbelieving it. On a less radical version, there are bodies of evidence to which rationality permits you to respond by believing a particular proposition or suspending on it, and there are bodies of evidence to which rationality permits you to respond by disbelieving a particular proposition or suspending on it, but there are none that permit all three attitudes. Perhaps, for instance, rationality permits me and each of my friends to believe that Biden will win, and also permits us to suspend judgment on it, but doesn't permit us to believe he will lose. That would not be radical. In contrast, perhaps my godson is rationally permitted to believe Joe will win, but also rationally permitted to believe he'll lose, and also rationally permitted to suspend judgment. That would be radical.

In the credal case, rationality might permit me to have any credence between 0.75 and 0.77 that Biden will win. That would not be radical. But it might permit me to have any credence between 0.1 and 0.9. That would be radical.

2.4 Common or rare?

Finally, whatever attitudes your permissivism applies to, whether it is interpersonal or intrapersonal, and whether it is radical or moderate, it might also be permissive often or rarely. According to common permissivism, most evidence permits a number of different attitudes; according to rare permissivism, most evidence does not.

A rare permissivist, for instance, might think that when evidence is very complex or very sparse, rationality is permissive, but also think that these situations are rather rare. For instance, our evidence is complex when we are trying to predict the behaviour of a complex system, like a society voting for a leader or the effect of an intervention in our climate system. And our evidence is sparse when we face an entirely new phenomenon about which we have little relevant prior evidence, such as the emergence of a novel viral threat or a new political player about whom we have almost no evidence. But usually, when we predict the effect of eating a type of food we've enjoyed many times before, or we set our credence that we'll see a goldfinch in the park we've walked in for ten years, our evidence is reasonably substantial, not terribly complex, and it largely points in the same direction.

A common permissivist, in contrast, thinks that rationality is permissive in the presence of nearly all evidence. Suppose, for instance, you think that the credences it is rationally permissible to have in response to a body of evidence are precisely the credences you could obtain by starting with a permissible prior credence function and updating on the evidence in the required way. You will then be permissivist about most posterior credences if you think that the priors that are permissible are varied enough and the evidence we typically have is sufficiently inconclusive to ensure that there are at least two different permissible priors that update on that evidence to give different posteriors.

2.5 My brand of permissivism

In this book, I'll defend two versions of permissivism. First, a radical common version of interpersonal permissivism about credences. Second, a slightly less radical and slightly less common version of intrapersonal permissivism about credences.

As I mentioned above, I will motivate these versions by arguing within epistemic utility theory; and I will appeal crucially to the notion of epistemic risk. Rationality permits a variety of attitudes to risk in the practical sphere. Faced with the same risky choice, you might be willing to gamble because you are risk-inclined, and I might be unwilling because I am risk-averse, but we are both rational and neither more rational than the other. On my Jamesian account, rationality also permits different attitudes to risk in the epistemic sphere. And different attitudes to epistemic risk warrant different credal attitudes prior to receiving any evidence. This delivers *inter*personal permissivism at least for the empty bodies of evidence to which we respond with our prior credences. What's more, because rationality permits a wide range of different attitudes to risk, it gives a *radical* version of interpersonal permissivism about priors—that is, there is a great variety of rationally permissible priors. And indeed it gives such a radical version that, for almost any evidence we subsequently receive, if we update each of this multitude of rationally permissible priors with that evidence, we get a multitude of rationally permissible posteriors. So it gives a *radical* and *common* version of interpersonal permissivism—rationality is radically permissive in the presence of most bodies of evidence. And finally, it turns out that even once an individual's risk attitudes have been fixed, it is often the case that many different priors and therefore many different posteriors are permitted. And that gives a radical common version of *intra*personal permissivism, but one that is slightly less radical and rather less common, because fixing the risk attitude does cut down the range of permissible priors and thus the range of permissible posteriors quite significantly.

In brief...

In this chapter, I built on and slightly extended Elizabeth Jackson's (2019) taxonomy of different species of permissivism about epistemic rationality, and I located within that framework the positions for which I'll argue in this essay. They are:

- Radical common interpersonal permissivism about credences.

  That is, for most bodies of evidence, the range of credal states for which there is some individual who might rationally respond to that evidence with that credal state is wide.

- Moderate common intrapersonal permissivism about credences.

  That is, for most bodies of evidence and many individuals, the range of credal states the individual might rationally adopt in response to that evidence is narrower, but it nonetheless contains more than one option.

3 Epistemic risk and epistemic utility

Let me begin by looking at a different attempt to argue for epistemic permissivism on the basis of epistemic risk. This is due to Thomas Kelly (2014), and also inspired by William James' (1897) 'The Will to Believe'. Kelly is concerned not with credences but with full beliefs, and, for them, I think his argument succeeds. However, as Sophie Horowitz (2017) points out, his argument cannot be transferred to the case of credences. That will lead us to seek a different approach, which we'll present in the remainder of the book.

3.1 William James' two duties and permissivism for full beliefs

Kelly begins by reminding us of William James' distinction between the two goals we have when we form beliefs—we want to believe truths, and we want not to believe falsehoods.

> There are two ways of looking at our duty in the matter of opinion,—ways entirely different, and yet ways about whose difference the theory of knowledge seems hitherto to have shown very little concern. We must know the truth; and we must avoid error,—these are our first and great commandments as would-be knowers; but they are not two ways of stating an identical commandment, they are two separable laws. [...]
>
> Believe truth! Shun error!—these, we see, are two materially different laws; and by choosing between them we may end by coloring differently our whole intellectual life. We may regard the chase for truth as paramount, and the avoidance of error as secondary; or we may, on the other hand, treat the avoidance of error as more imperative, and let truth take its chance. [...]
>
> We must remember that these feelings of our duty about either truth or error are in any case only expressions of our passional life. (James, 1897, 18)

When we have a belief, it gives us a chance of being right, but it also runs the risk of being wrong. In contrast, when we withhold judgment on a proposition, we run no risk of being wrong, but we give ourselves no chance of being right. Pursuing the two Jamesian goals, then, pulls you in two directions. You could guarantee that you achieve the first by believing every proposition, since you'd thereby believe all truths. And you could guarantee you achieve the second by believing no propositions, since you'd thereby believe no falsehoods. Where you should fall between these two extremes, Kelly says, depends on your attitudes to epistemic risk. The epistemically risk-inclined will put more weight on the first of James' two aspects of our duty in the matter of opinion—*Believe truth!* The epistemically risk-averse will put more weight on the second—*Shun error!*

To make this precise, we might adopt the epistemic utility theory framework described in the Introduction, and encode your attitudes to epistemic risk in your epistemic utilities. The idea is that, just as you assign different pragmatic value to your actions depending on how the world is, and measure that by your pragmatic utility for that action given the world is that way, so you might assign different purely epistemic value to your doxastic attitudes and states depending on how the world is, and measure that by your epistemic utility for those attitudes and states given the world is that way. According to Kelly's reading of James, the weights you place on the two different duties the matter of opinion are best encoded in your epistemic utilities. The greater the weight you put on *Believe truth!*, the more highly you value true beliefs and false disbeliefs; the greater the weight you put on *Shun error!*, the more highly you disvalue false beliefs and true disbeliefs.

We can make this precise by appealing to a thesis introduced independently by Kenny Easwaran and Kevin Dorst, which finds its inspiration in a suggestion of Carl Hempel's (Hempel, 1962; Easwaran, 2016; Dorst, 2019). Suppose:[8]

(i) I assign a positive epistemic utility of $R > 0$ to believing a truth or disbelieving a falsehood—this is my utility for getting things right;

(ii) I assign a negative epistemic utility (or positive epistemic disutility) of $-W < 0$ to believing a falsehood or disbelieving a truth—this is my utility for getting things wrong; and

(iii) I assign a neutral epistemic utility of 0 to withholding judgment.

Then, the more averse you are to epistemic risk, the greater weight you will put on *Shun error!* and the less you will put on *Believe truth!*, and so the greater will be $\mathcal{W}$ in comparison with R. That is, you will dislike being wrong much more than you like being right. The more favourable you are to epistemic risk, the greater weight you put on *Believe truth!* and the less on *Shun error!*, and so the greater will be R in comparison with $\mathcal{W}$. That is, you will like being right more than you dislike being wrong.

The James-Kelly-Easwaran-Dorst account of epistemic utility just sketched assumes veritism, the axiological view that we met briefly in the Introduction and which says that the sole fundamental source of epistemic utility for a doxastic attitude is its accuracy—the closer the attitude comes to representing the world accurately, the greater the utility. So the epistemic utility

[8]What I present here is the simplest version of Easwaran's and Dorst's thesis, but it suffices to run Kelly's argument.

of a true belief is greater than the epistemic utility of a false belief; and withholding judgment lies between the two. But, according to James and Kelly, veritism leaves open a wide range of rational permissible ways of valuing accuracy. Following James, Kelly takes your attitudes to epistemic risk, encoded in your epistemic utilities R and $\mathcal{W}$, to be "expressions of your passional life". That is, he takes them to be like your tastes and values—attitudes that are not governed by some external standard, but which are instead, to a great extent, up to you. That is, there is a large permissible range of values for R and $\mathcal{W}$ from which you might choose.

This gives us a way of encoding your attitudes to epistemic risk within your epistemic utilities in the case of full beliefs. And we can now use that to derive permissivism about rational belief from the Jamesian permissivism about attitudes to epistemic risk. Suppose that there is some way to measure, for each proposition, how likely my evidence makes that proposition. That is, suppose there is a unique evidential probability function of the sort that J. M. Keynes and Timothy Williamson envisage (Keynes, 1921; Williamson, 2000). Then, if r is how likely my evidence makes the proposition X, then we can calculate the expected epistemic utility of having a belief in X, the expected epistemic utility of having a disbelief in X, and the expected epistemic utility of withholding judgment on X:

(i) the expected epistemic utility of believing X is $rR + (1-r)(-W)$,

(ii) the expected epistemic utility of disbelieving X is $r(-W) + (1-r)R$, and

(iii) the expected epistemic utility of withholding judgment is 0.

Having done so, we are rationally required to pick an attitude that has maximal expected epistemic utility. Here are the crucial facts:

$$\begin{aligned}
\text{Believe } X \geq \text{Disbelieve } X \quad &\Leftrightarrow \quad r \geq \frac{1}{2} \\
\text{Believe } X \geq \text{Withhold on } X \quad &\Leftrightarrow \quad r \geq \frac{W}{R+W} \\
\text{Disbelieve } X \geq \text{Withhold on } X \quad &\Leftrightarrow \quad r \leq \frac{R}{R+W}
\end{aligned}$$

And here are the consequences. There are three cases: risk-averse, risk-neutral, and risk-inclined.

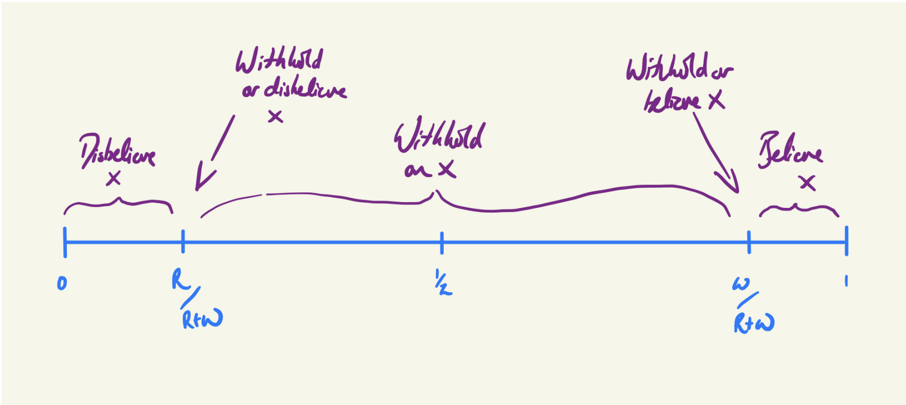

Figure 1: The risk-averse case (i.e., $\mathcal{W} > R$): the attitudes required or permitted by different evidential probabilities.

First, risk-averse. Suppose $\mathcal{W} > R$, so that $\frac{R}{R+W} < \frac{1}{2} < \frac{W}{R+W}$. Then

$$\begin{aligned}
\frac{1}{2} < \frac{W}{R+W} < r &\Rightarrow \text{Believe } X \\
\frac{1}{2} < \frac{W}{R+W} = r &\Rightarrow \text{Believe } X \text{ or Withhold on } X \\
\frac{R}{W+R} < r < \frac{W}{R+W} &\Rightarrow \text{Withhold on } X \\
r = \frac{R}{R+W} < \frac{1}{2} &\Rightarrow \text{Disbelieve } X \text{ or Withhold on } X \\
r < \frac{R}{R+W} < \frac{1}{2} &\Rightarrow \text{Disbelieve } X
\end{aligned}$$

This is illustrated in Figure 1.

Second, risk-neutral or risk-inclined. Suppose $\mathcal{W} \leq R$, so that $\frac{W}{R+W} \leq \frac{1}{2} \leq \frac{R}{R+W}$. Then

$$\begin{aligned}
\frac{1}{2} < r &\Rightarrow \text{Believe } X \\
r = \frac{1}{2} &\Rightarrow \text{Believe } X \text{ or Withhold on } X \text{ or Disbelieve } X \\
r < \frac{1}{2} &\Rightarrow \text{Disbelieve } X
\end{aligned}$$

This is illustrated in Figure 2.

So, perhaps unsurprisingly, the more you disvalue having an attitude that is wrong—a false belief or a true disbelief—and the less you value having an attitude that is right—a true belief or a false disbelief—the stronger the evidence in favour of a proposition will have to be in order to make it

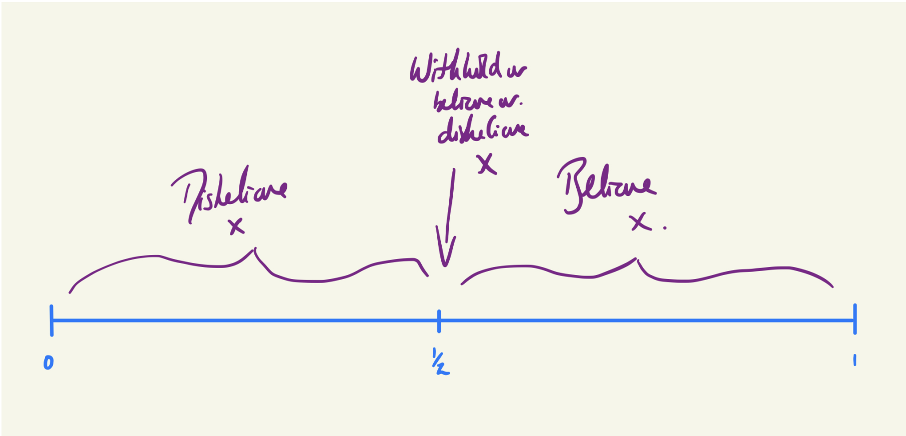

Figure 2: The risk-neutral and risk-inclined case (i.e. $R \geq W$): the attitudes required or permitted by different evidential probabilities.

rational to believe and the stronger the evidence against will have to be in order to make it rational to disbelieve. For the risk-averse agent, $\mathcal{W} > R$, and the more risk-averse, the greater the ratio of $\mathcal{W}$ to R. So the more risk-averse you are, the stronger the evidence in favour of a proposition has to be to compel you to believe it rather than withhold judgment on it, and the stronger the evidence against it must be to compel you to disbelieve rather than withhold judgment. For the risk-inclined agent, $\mathcal{W} < R$, and then it doesn't matter how risk-inclined you are: the threshold for belief is $\frac{1}{2}$ regardless. And for the risk-neutral agent, it is the same.

Now, as we noted above, Kelly follows William James in assuming that there are many different values of R and $\mathcal{W}$ that are rationally permissible. It is permissible to disvalue believing falsehoods a lot more than you value believing truths; it is permissible to disvalue that just a lot less; and many attitudes in between are also permissible. If that is the case, different individuals might have the same evidence while rationality requires of them different doxastic attitudes. Rationality might require of one of them that they believe the proposition, because they disvalue being wrong only a little more than they value being right, and their evidence makes the proposition likely enough that it is worth taking the risk. But it might require of the other individual that they suspend judgment, because the ratio between their disvalue for false belief and value for true belief is much greater. In our reconstruction of Kelly's thinking, using Easwaran's and Dorst's framework, we identify the values you pick for R and $\mathcal{W}$ with your attitudes to epistemic risk. So different doxastic attitudes are permissible in the face of the same evidence because different attitudes to epistemic risk are permissible and those attitudes to epistemic risk play a key role in determining what it is rational for you to believe.

Now, there are a number of features of Kelly's approach that are worth noting.

First, it belongs to epistemic utility theory. As we noted in the Introduction, this says that we determine the epistemic rationality of an agent's doxastic state as follows: we treat the agent as facing a decision problem in which the available options are doxastic states; we take the utility of those doxastic states to be their purely epistemic utility; we apply the appropriate decision rule to that decision problem with those utilities; and whatever states it deems rationally permissible are the ones that epistemic rationality permits. For Kelly, when we consider whether a doxastic attitude towards X is rational or not, the available options are: *Believe X*, *Disbelieve X*, *Withhold judgment on X*. The utility of each option given each way the world might be is determined by the agent's epistemic utilities, R and $\mathcal{W}$. And the correct decision rule says: an option is rational for an agent just in case it maximises expected utility from the point of view of the evidential probabilities relative to that agent's total evidence. We will meet other versions of epistemic utility theory below. Indeed, epistemic utility theory is the approach taken in this book. But the options I consider after this chapter will be individual credences or whole credal states, represented by credence functions; the utilities will be determined by so-called scoring rules in the way we set out in the next section; and the decision rules will be quite different—for instance, they won't involve evidential probabilities.

Second, note that Kelly manages to show that epistemic rationality might be permissive even if there is a unique evidential probability measure—that's quite a feat! So even if you think you can solve the problem of which precise probability is demanded by the very sparse evidence and the very complex evidence we described in the Introduction, still you should countenance a form of epistemic permissivism if you agree with James and Kelly that there are different permissible values for R and $\mathcal{W}$. Thus, Kelly's approach is markedly different from those where we claim that permissivism results from a sort of indeterminacy in the measure of evidential support. One consequence of this is that Kelly's permissivism is much more widespread—it does not only arise when evidence is very sparse or very complex. It arises even when our evidence is unequivocal though still uncertain, such as when you know for certain that there are seven purple balls and three green balls in the urn before you and you're about to pick one in a way that you know is completely random. For Kelly, even though the evidential probability that you'll draw a purple ball is $\frac{7}{10}$, some individuals will be required to believe that they'll draw a purple ball—those for whom $\frac{W}{R+W} < \frac{7}{10}$—while other, more risk-averse individuals will be required to withhold judgment—those for whom $\frac{W}{R+W} > \frac{7}{10}$.

Just how common permissivism is on Kelly's account depends on which combinations of R and $\mathcal{W}$ you consider rational. We might think that

$\mathcal{W} = 3$ and $R = 2$ is permitted, but $\mathcal{W} = 1,000$ and $R = 1$ is not. William James himself perhaps suggests this in his response to Clifford, whom he thinks is unreasonably risk-averse in the matter of belief:

> [H]e who says, "Better go without belief forever than believe a lie!" merely shows his own preponderant private horror of becoming a dupe. (James, 1897)

Though it's not clear that he intends this as a charge of irrationality. I won't try to adjudicate this matter here. It will arise again in the context of credal epistemology in Section 5.2 below.

Third, it might seem at first that Kelly's argument gives interpersonal permissivism at most. After all, for a fixed attitude to epistemic risk, encoded in R and $\mathcal{W}$, and a unique evidential probability for X given your evidence, let's say r, it might seem that there is always a single attitude—belief in X, disbelief in X, or judgment withheld about X—that maximises expected epistemic value. But this isn't always true, as we can see from the lists and figures above. After all, if $r = \frac{R}{R+W}$, then it turns out that disbelieving and withholding have the same expected epistemic value, and if $r = \frac{W}{R+W}$, then believing and withholding have the same expected epistemic value. And in those cases, it would be rationally permissible for an individual to pick either.

Fourth, and relatedly, it might seem that Kelly's argument gives only narrow permissivism, since it allows for cases in which believing and withholding are both rational, and it allows for cases in which disbelieving and withholding are both rational, but it doesn't allow for cases in which all three are rational. But that again is a mistake. If you value believing truths exactly as much as you value believing falsehoods, so that $R = W$, and if the objective evidential probability of X given your evidence is $r = \frac{1}{2}$, then believing, disbelieving, and withholding judgment are all permissible. Having said that, there is some reason to say that it is not rationally permissible to set $R = W$. After all, if you do, and if $r = \frac{1}{2}$, then it is permissible to both believe X and believe the negation of X at the same time, and that seems wrong.

So Kelly's argument from epistemic risk to permissivism about full beliefs, supplemented with Easwaran's and Dorst's precise account of veritist epistemic value for such doxastic states, delivers a surprisingly strong form of permissivism. Indeed, as strong as we hope to establish for credences. Unfortunately, however, his approach won't work for credences, as we'll see in the next section.

3.2 Encoding epistemic risk in epistemic utilities for credences

Kelly's account of the epistemic value of beliefs and disbeliefs—inspired by William James and made precise by Kenny Easwaran and Kevin Dorst—

delivers permissivism about beliefs because there are propositions X, bodies of evidence we might have that give evidential probability r to X, and different legitimate ways of valuing getting things right and wrong—namely, R_1 and $-W_1$, respectively, on the one hand, and R_2 and $-W_2$ on the other—such that r assigns highest expected epistemic value to believing X when epistemic value is determined by R_1 and $-W_1$, but assigns highest expected epistemic value to withholding judgment on X when that value is determined by R_2 and $-W_2$. If we now try to extend this approach to the credal case, we face the problem that, for any X, any evidential probability r that our evidence might bestow on X, and any legitimate way of measuring the epistemic value of credences, r assigns highest expected epistemic value to having credence r in X. That is, given any evidential probability, relative to all legitimate ways of measuring the epistemic value of credences, that probability assigns highest expected epistemic value to the same single credence, namely, the one that matches the evidential probability. So we do not obtain permissivism about credences from permissivism about measures of epistemic value in the same way that Kelly did in the case of full belief. This was first pointed out by Sophie Horowitz (2017).

The reason is that all legitimate measures of the epistemic value of credences are *strictly proper*. Let's spell out what this means. The James-Kelly-Easwaran-Dorst measures of epistemic utility for beliefs apply to individual doxastic attitudes; in particular, individual beliefs. They say, for a particular belief, how valuable it is if the proposition believed is true and if the proposition believed is false. The analogue to this in the credal case is a *scoring rule*. This is a pair of functions, $\mathfrak{s}(0, -)$ and $\mathfrak{s}(1, -)$: $\mathfrak{s}(0, x)$ measures the epistemic value of having credence x in a false proposition; $\mathfrak{s}(1, x)$ measures the epistemic value of having credence x in a true proposition.[9] Some examples:

(i) *The absolute value scoring rule*

- $\mathfrak{a}(0, x) = -x$
- $\mathfrak{a}(1, x) = -(1 - x)$

(ii) *The quadratic scoring rule*

[9]Scoring rules were originally used not to measure the epistemic value of a credence, but to specify some pragmatic reward you might receive for having a particular credence in a certain situation. For instance, if you run a weather forecasting firm, you might wish to reward your employees differently for the accuracy of their forecasts. If it does indeed rain by noon on Monday, you might wish to give a higher bonus to one employee, who posted 80% confidence that it would, than another, who posted 20% confidence. In fact, as they were originally discussed, the term 'scoring rule' covered ways of penalising credences, rather than rewarding them. But a penalty function is just the negative of a reward function, so we keep the terminology here. Brier (1950) describes one of the first scoring rules, which was indeed used in the meteorological setting; Savage (1971) proves some of the fundamental properties of the species of scoring rule that we'll consider here.

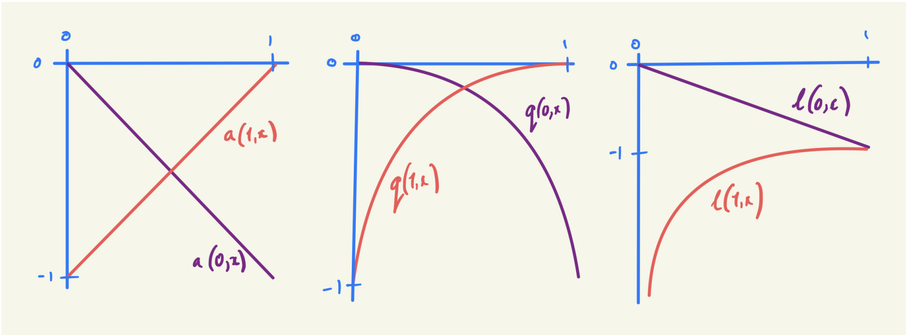

Figure 3: Three scoring rules, $\mathfrak{a}$, $\mathfrak{q}$, $\mathfrak{l}$.

- $\mathfrak{q}(0, x) = -x^2$
- $\mathfrak{q}(1, x) = -(1 - x)^2$

(iii) *The enhanced log scoring rule*

- $\mathfrak{l}(0, x) = -x$
- $\mathfrak{l}(1, x) = -(x - \log x)$

Given an evidential probability r and a credence x, the expected epistemic utility that r assigns to x is

$$r\mathfrak{s}(1, x) + (1 - r)\mathfrak{s}(0, x)$$

Now, we say that the scoring rule $\mathfrak{s}$ is *strictly proper* if

(i) $\mathfrak{s}(1, p)$ and $\mathfrak{s}(0, p)$ are continuous functions of p, and

(ii) for any $0 \leq p \neq q \leq 1$,

$$p\mathfrak{s}(1, p) + (1 - p)\mathfrak{s}(0, p) > p\mathfrak{s}(1, q) + (1 - p)\mathfrak{s}(0, q)$$

That is, $\mathfrak{s}$ is strictly proper if, for each probability $0 \leq p \leq 1$, the credence p expects itself to be best, epistemically speaking. That is, it expects every other credence to be worse than it expects itself to be. Recalling the examples above: $\mathfrak{q}$ and $\mathfrak{l}$ are strictly proper, while $\mathfrak{a}$ is not.[10]

So, if we follow Kelly's Jamesian lead and encode our attitudes to epistemic risk in our epistemic utilities, and if only strictly proper scoring rules

[10] We briefly sketch the proofs:

- The quadratic score:

$$\frac{d}{dx}\left(r\mathfrak{q}(1, x) + (1 - r)\mathfrak{q}(0, x)\right) = -\frac{d}{dx}\left(r(1 - x)^2 + (1 - r)x^2\right) = 2(r - x) = 0$$

iff $x = r$. So the expected value of $\mathfrak{q}$ relative to r is maximised uniquely at $x = r$.

can be legitimate measures of the epistemic value of a credence, and if we retain the requirement that we choose only options that maximise expected utility from the point of view of the evidential probabilities, we will not obtain permissivism about credences. After all, if r is the evidential probability of X, then assigning credence r to X uniquely maximises expected epistemic utility. It is thus the unique rational response to that evidence.

Scoring rules are measures of the epistemic value of having an individual credence. But we'll also be interested in measures of the epistemic value of an entire credal state: that is, a collection of credences, each assigned to a different proposition. Thus, we are interested not only in the epistemic value of my credence 0.9 that it will rain in the next hour, but in the epistemic value of my whole credal state, which includes that credence in imminent rain, as well as my credence of 0.7 that it will be sunny tomorrow, my credence that Joe Biden will be inaugurated on 20th January, my credence that I have had an asymptomatic case of COVID-19, my credence that the water in my tap is safe to drink, and so on. We represent this credal state by a *credence function*. The domain of this function is the set of all propositions to which I assign a credence—call this my *opinion set* or *agenda* and denote it $\mathcal{F}$. Then my credence function C takes each proposition X in $\mathcal{F}$ and returns my credence $C(X)$ in X. For any X in $\mathcal{F}$, $C(X)$ is at least 0 and at most 1—that is, $0 \leq C(X) \leq 1$.

A measure of epistemic value, or an epistemic utility function, is then a function $\mathfrak{EU}$ that takes

(i) a credence function C defined on a set of propositions $\mathcal{F}$, and

(ii) a possible world, specified as a classically consistent assignment of truth values to the propositions in $\mathcal{F}$,

and returns

(iii) a real number or $-\infty$ that measures the epistemic value of having that credence function at that world.

- The enhanced log score:

$$\frac{d}{dx}\left(r\mathfrak{l}(1,x) + (1-r)\mathfrak{l}(0,x)\right) = -\frac{d}{dx}\left(r(x - \log x) + (1-r)x\right) = \frac{r}{x} - 1 = 0$$

iff $x = r$. So the expected value of $\mathfrak{l}$ relative to r is maximised uniquely at $x = r$.

- The absolute value score:

$$r\mathfrak{a}(1,x) + (1-r)\mathfrak{a}(0,x) = -(r(1-x) + (1-r)x) = (2r-1)x - r$$

is a straight line. It is increasing if $r > \frac{1}{2}$, and thus maximised at $x = 1$; it is decreasing if $r < \frac{1}{2}$, and thus maximised at $x = 0$; it is constant if $r = \frac{1}{2}$, and thus maximised for any $0 \leq x \leq 1$.

We'll assume throughout that $\mathcal{F}$ is a finite algebra. How the proposal of this essay might be extended to the infinite case is an interesting question. But there will be plenty to deal with in the finite case, so I restrict attention to that here.

Whenever we talk of a finite algebra $\mathcal{F}$, we'll talk of $\mathcal{W}$ as the set of possible worlds relative to that algebra. That is, $\mathcal{W}$ is the set of classically consistent assignments of truth values to the propositions in $\mathcal{F}$. Since $\mathcal{F}$ is a finite algebra, to each world w in $\mathcal{W}$ there corresponds a proposition in $\mathcal{F}$ that is true at that world only and not at any others. We will abuse notation and denote both the world and this proposition $\mathcal{W}$.

A natural way to generate measures of the epistemic value of a credal state from measures of the epistemic value of the individual credences that compose it is to take the former to be the sum of the latter. If $\mathfrak{s}$ is a scoring rule, let

$$\mathfrak{EU}_{\mathfrak{s}}(C, w) = \sum_{X \in \mathcal{F}} \mathfrak{s}(w(X), C(X))$$

where we abuse notation still further and, for any possible world w in $\mathcal{W}$, we write w for the following function: $w(X) = 1$ if X is true at w, and $w(X) = 0$ if X is false at w.[11] This gives the following measures using our examples from above:

(i) *The absolute value measure*

$$\mathfrak{A}(C, W) := \mathfrak{EU}_{\mathfrak{a}}(C, w) = \sum_{X \in \mathcal{F}} \mathfrak{a}(w(X), C(X)) = -\sum_{X \in \mathcal{F}} |w(X) - C(X)|$$

(ii) *The Brier score*

$$\mathfrak{B}(C, w) := \mathfrak{EU}_{\mathfrak{q}}(C, w) = \sum_{X \in \mathcal{F}} \mathfrak{q}(w(X), C(X)) = -\sum_{X \in \mathcal{F}} |w(X) - C(X)|^2$$

(iii) *The enhanced log score*

$$\mathfrak{L}(C, w) := \mathfrak{EU}_{\mathfrak{l}}(C, w) = \sum_{X \in \mathcal{F}} \mathfrak{l}(w(X), C(X)) =$$
$$-\left(\sum_{X \in \mathcal{F}} C(X) - \sum_{X \in \mathcal{F}} w(X) \log C(X)\right)$$

[11]While it might be natural, this is by no means the only way to generate a measure of the epistemic utility of a whole credence function from measures of the epistemic utility of the credences it assigns. For instance, we might include in the summation only some subset of the credences it assigns. Or we might weight each credence's score before summing, perhaps giving different weights to credences in stronger propositions. Or we might score each credence using a different scoring rule. And indeed we might not use summation at all, nor even take the epistemic utility of the whole credal state to be determined by the epistemic utility of the individual credences. But of course any greater permissivism about measures of epistemic utility can only give greater permissivism about rational credence. And so here we will continue to measure epistemic utility in the more restricted way just described. Thanks to Teddy Seidenfeld for discussion of this.

Some credence functions on a finite algebra $\mathcal{F}$ are probability functions, and some are not. A credence function C on $\mathcal{F}$ is a *probability function* or a *probabilistic credence function* if

(P1) $\sum_{w \in \mathcal{W}} C(w) = 1$;

(P2) for any proposition X in $\mathcal{F}$, $C(X) = \sum_{w \in X} C(w)$.

That is, the sum of the credences that C assigns to each of the worlds in $\mathcal{W}$ is 1, and the credence that C assigns to a proposition X is the sum of the credences that C assigns to the worlds at which X is true.

We say that an epistemic value measure $\mathfrak{EU}$ is *strictly proper* if, for any probability function P defined on $\mathcal{F}$, and any other credence function C defined on $\mathcal{F}$,

$$\sum_{w \in \mathcal{W}} P(w)\mathfrak{EU}(P, w) > \sum_{w \in \mathcal{W}} P(w)\mathfrak{EU}(C, w)$$

Note that $\mathfrak{s}$ is strictly proper iff $\mathfrak{EU}_{\mathfrak{s}}$ is strictly proper. So $\mathfrak{B}$ and $\mathfrak{L}$ are strictly proper, while $\mathfrak{A}$ is not. Now, again, we have that, if R is the evidential probability function, and $\mathfrak{EU}$ is strictly proper, R will expect the credence function R itself to be best. Again, Kelly's route to permissivism leads to a dead end in the credal case.

Why do we say that only strictly proper epistemic utility measures can be legitimate? There are a number of arguments for this. I'll describe three. I won't spend a great deal of time on these. I am persuaded by the second, and possibly the third; here, I will simply assume that they establish that no scoring rule that isn't strictly proper can be used to measure epistemic utility, and possibly something stronger than that. If this is wrong, then it opens up the possibility that we might find an argument for permissivism that proceeds like Kelly's: we encode different permissible attitudes to epistemic risk in different legitimate epistemic utility functions, which now include more than just the strictly proper ones; and we show that some of the different epistemic utility functions corresponding to different attitudes to epistemic risk render different credences functions rationally permissible in the presence of a fixed body of evidence. So you might read this book as investigating how epistemic risk might yield epistemic permissivism, even if Kelly's approach won't work.

Our first argument for strict propriety comes originally from Graham Oddie, though it also forms the cornerstone of Jim Joyce's second attempt at characterising epistemic utility functions (Oddie, 1997; Joyce, 2009).[12] The argument begins with the observation that, though the permissivist claims that there are evidential situations in which rationality permits more than one credal response, they also admit that there are evidential situations in

[12]See Hájek (2008), Konek (ta), and Blackwell & Drucker (2019) for objections.

which rationality demands a single credal response. For instance, most permissivists will say that, if you learn the objective chances of the propositions about which you have an opinion, then the only rational response is to adopt those chances as your credences. Now, Oddie and Joyce say, for any probability function P on $\mathcal{F}$, you might learn that P gives the objective chances. And if you were to learn that, rationality would require that you adopt P as your credence function. Now, suppose that P expects some other credence function $C \neq P$ to be at least as good as it expects itself to be. That is,

$$\sum_{w \in \mathcal{W}} P(w)\mathfrak{EU}(P, w) \leq \sum_{w \in \mathcal{W}} P(w)\mathfrak{EU}(C, w)$$

Then it seems that rationality must permit you to move from P to C without any further evidence—after all, by your own lights, doing so is not expected to lose you any epistemic value. But that's precisely what we all agree is not true—we agree that rationality forbids you from moving away from P while your evidence is still only that P gives the objective chances. Thus, we must have

$$\sum_{w \in \mathcal{W}} P(w)\mathfrak{EU}(P, w) > \sum_{w \in \mathcal{W}} P(w)\mathfrak{EU}(C, w)$$

That is, $\mathfrak{EU}$ must be strictly proper.

My own argument, as well as D'Agostino and Sinigaglia's, is based on an axiomatic characterization of measures of epistemic value. Both appeal to properties other than strict propriety, but each ends up characterising precisely the strictly proper epistemic utility functions (Pettigrew, 2016a; D'Agostino & Sinigaglia, 2010). Both begin with the idea that a credal state is more valuable, epistemically speaking, the better it corresponds to reality. This is the thesis I called *veritism* in the Introduction. To make this veritist assumption precise, we introduce two pieces of machinery. First, the ideal credence function at a possible world. This is the one that represented the world with maximal accuracy. Both arguments claim that the ideal credence function at a world is the one that gives maximal credence to all truths and minimal credence to all falsehoods. That is, if for any world w in $\mathcal{W}$, we abuse notation as before and define the function w as above—so that $w(X) = 1$ if X is true at w and $w(X) = 0$ if X is false at w—then w is the ideal credence function at world w. The second piece of machinery is a measure of distance $\mathfrak{D}$ from one credence function to another. The idea is that the epistemic value of a credence function C at a world w is given by how close C lies to the ideal credence function at that world, namely, w. That is, $\mathfrak{EU}(C, w) = -\mathfrak{D}(w, C)$. Both arguments then proceed by placing conditions on $\mathfrak{D}$. If $\mathfrak{D}$ satisfies all of my conditions, and if $\mathfrak{EU}(C, w) = -\mathfrak{D}(w, C)$, then $\mathfrak{EU}$ is a strictly proper epistemic value measure generated by a strictly proper scoring rule. And if $\mathfrak{D}$ satisfies all of

D'Agostino and Sinigaglia's condition, and if $\mathfrak{EU}(C, w) = -\mathfrak{D}(w, C)$, then $\mathfrak{EU}$ is the Brier score $\mathfrak{B}$, which is strictly proper, as we saw above.

3.3 Foundational results in epistemic utility theory

Before moving on, it's worth noting a handful of facts about strictly proper scoring rules and the strictly proper epistemic utility functions that they generate. We will have cause to refer back to these a number of times in what follows. Together, these provide the foundational results in the credal version of epistemic utility theory, and they ground the epistemic utility arguments for the following norms that we typically take to govern our credences: Probabilism, the Principal Principle, and two synchronic versions of Conditionalization, the Bayesian's favoured updating rule.[13]

3.3.1 Probabilism

The first norm is Probabilism, which says that your credences should satisfy the axioms of the probability calculus. We usually state those axioms as follows:

(P1′) $C(\top) = 1$ and $C(\bot) = 0$

That is, you should have maximal credence in a tautology, and minimal credence in a contradiction.

(P2′) $C(X \vee Y) = C(X) + C(Y)$ if X and Y are mutually exclusive.

That is, your credence in a disjunction of two mutually exclusive propositions should be the sum of your credences in the disjuncts.

Above, we stated the axioms differently:

(P1) $\sum_{w \in \mathcal{W}} C(w) = 1$;

(P2) $\sum_{w \in X} C(w) = C(X)$.

It is easy to see that these are equivalent: (P1) + (P2) $\Leftrightarrow$ (P1′) + (P2′).

The norm can then be stated more carefully:

> **Probabilism** At any time in your epistemic life, it is rationally required that, if C is your credence function at that time, then C is a probability function.

To establish this, epistemic utility theory appeals to the following result (Joyce, 2009; Predd et al., 2009):

[13] We'll leave an argument for a diachronic version of Conditionalization until later (Section 6.3).

Theorem 1 *Suppose $\mathfrak{EU}$ is a strictly proper epistemic utility function, and suppose C is a credence function defined on $\mathcal{F}$. Then:*

(I) *If C is not a probability function, then there is an alternative credence function $C^\star$ defined on $\mathcal{F}$ that is a probability function such that, for all possible worlds w in $\mathcal{W}$,*

$$\mathfrak{EU}(C, w) < \mathfrak{EU}(C^\star, w)$$

(II) *If C is a probability function, then there is no alternative credence function $C^\star \neq C$ defined on $\mathcal{F}$ such that, for all possible worlds w in $\mathcal{W}$,*

$$\mathfrak{EU}(C, w) \leq \mathfrak{EU}(C^\star, w)$$

In the language of decision theory: if you violate Probabilism, then your credence function is strictly dominated; if you satisfy it, it is not even weakly dominated. As I noted above, one of the key claims of epistemic utility theory is that an account of epistemic utility together with an appropriate decision rule will deliver a verdict about what is rational. In this case, the decision rule is one that I have called Undominated Dominance: it says that it is irrational to choose an option that is strictly dominated by some alternative option that isn't itself even weakly dominated (Pettigrew, 2016a). From this, together with the claim that epistemic utility functions must be strictly proper, we conclude Probabilism.

3.3.2 The Principal Principle

The second norm is the Principal Principle. At the beginning of the paper in which he first discusses this norm, and in which he gives it this name, David Lewis gestures informally towards a number of versions (Lewis, 1980). Here's one: rationality requires that, if your total evidence entails that ch is the probability function that gives the objective chances, and if your credence function is C, then $C(X) = ch(X)$, for all X in $\mathcal{F}$. Here's another: rationality requires that, if C is your prior credence function at the beginning of your epistemic life, and C_{ch} is the proposition that says that ch is the probability function that gives the current objective chances, and $C(C_{ch}) > 0$, then $C(X|C_{ch}) = ch(X)$. And here's a third: rationality requires that if $\mathcal{C}$ is the set of epistemically possible objective chance functions, then C should be a weighted average of the members of $\mathcal{C}$.

The version I will state is closest to the latter. As we see, it appeals to the notion of an epistemically possible objective chance function. A probability function ch is an epistemically possible objective chance function at a particular time if your total evidence at that time is compatible with ch giving the objective chances; that is, if your evidence does not rule out ch giving the objective chances. Then:

Principal Principle At any time in your epistemic life, if $\mathcal{C}$ is the set of epistemically possible objective chance functions for you at that time, then rationality requires that, if C is your credence function at that time, then C is a member of the closed convex hull of $\mathcal{C}$.

This uses a technical notion, namely, the notion of a *closed convex hull*. Given a set $\mathcal{X}$ of probability functions, the closed convex hull of $\mathcal{X}$ is the smallest set with the following properties: (i) it contains all the probability functions in $\mathcal{X}$; (ii) whenever it contains two probability functions, it contains all mixtures of them; (iii) whenever it contains an infinite sequence of probability functions that approach another probability function in the limit, it also contains that other probability function.[14]

This sounds like an overly technical requirement, but in straightforward cases it makes straightforward and intuitively plausible demands. Suppose, for instance, that I have been told that the urn before me either contains one blue ball and three orange ones or one orange ball and three blue ones. Then there are just two epistemically possible objective chance functions. The first: $ch_1(Blue) = \frac{1}{4}$, where *Blue* is the proposition that a ball drawn at random from the urn will be blue. The second: $ch_2(Blue) = \frac{3}{4}$. Then the Principal Principle demands that my current credence function is some mixture of these two. That is, there is $0 \leq \alpha \leq 1$ such that, for all X in $\mathcal{F}$, $C(X) = \alpha ch_1(X) + (1-\alpha) ch_2(X)$.

To establish this version of the Principal Principle, epistemic utility theory appeals to the following result (Pettigrew, 2013):

Theorem 2 *Suppose $\mathfrak{EU}$ is a strictly proper epistemic utility measure, $\mathcal{C}$ is a set of probability functions, and suppose C is a credence function defined on $\mathcal{F}$. Then:*

(I) *If C is not in the closed convex hull of $\mathcal{C}$, then there is an alternative credence function $C^\star$ defined on $\mathcal{F}$ that is in the closed convex hull of $\mathcal{C}$ such that, for all ch in $\mathcal{C}$,*

$$\sum_{w \in \mathcal{W}} ch(w) \mathfrak{EU}(C, w) < \sum_{w \in \mathcal{W}} ch(w) \mathfrak{EU}(C^\star, w)$$

(II) *If C is in the closed convex hull of $\mathcal{C}$, then there is no alternative credence function $C^\star \neq C$ defined on $\mathcal{F}$ such that, for all ch in $\mathcal{C}$,*

$$\sum_{w \in \mathcal{W}} ch(w) \mathfrak{EU}(C, w) \leq \sum_{w \in \mathcal{W}} ch(w) \mathfrak{EU}(C^\star, w)$$

Say that your credence function is *strictly chance dominated* by an alternative if every epistemically possible chance function expects the alternative to be

[14] If C and C' are probability functions, $C^\star$ is a mixture of C and C' if there is $0 \leq \lambda \leq 1$ such that $C^\star = \lambda C + (1-\lambda)C'$.

strictly better than yours; and say that it is *weakly chance dominated* by the alternative if they all expect it to be at least as good as yours, and some expect it to be better. So, if your credence function violates the Principal Principle, it is strictly chance dominated; and if it satisfies the Principal Principle, it is not. To establish the Principal Principle from this, epistemic utility theory appeals to the following decision rule, which we might call Undominated Chance Dominance: it says that it is irrational to choose an option that is strictly chance dominated by some alternative option that isn't itself even weakly chance dominated (Pettigrew, 2016a). From this, together with the claim that epistemic utility functions must be strictly proper, we establish the Principal Principle.

3.3.3 Plan Conditionalization

Finally, our third and fourth norms are slightly different formulations of the synchronic version of Conditionalization. It might seem that this is an oxymoron. Isn't the norm of Conditionalization quintessentially diachronic? Doesn't it concern how we update our credences from one time to another when we learn new evidence between those two times? Yes, it's true that this is how we usually understand Conditionalization, and indeed we'll argue for exactly such a diachronic principle in Section 6.3 below, building on an argument that Hannes Leitgeb and I once gave, and which Dmitri Gallow has recently improved significantly (Leitgeb & Pettigrew, 2010b; Gallow, 2019). But here we mention two arguments for closely related versions of a synchronic norm. This norm governs not the relationship between your credence functions at an earlier and at a later time, but between your credences at an earlier time and your plans at that very same time for how you will update those credences at the later time in the light of evidence that you'll receive in between. This is the sense in which it is synchronic: it governs the relationship between credence functions and updating plans, both of which are held at the same time.

An *updating plan* is a function that maps each body of total evidence you might have at the later time onto the new credence function that the plan says you should adopt if you do receive that evidence. Suppose $\mathcal{E} = \{E_1, \ldots, E_m\}$ is the set of propositions that give the different bodies of total evidence you might have at the later time. That is, by the later time, you might come to have total evidence E_1, or you might come to have total evidence E_2, and so on. Then a plan R takes each possible body of evidence E_i and returns a credence function R_i, which is the posterior you plan to have after acquiring E_i as your total evidence.

> **Plan Conditionalization (narrow scope)** If C is your credence function at the earlier time, and if you know that your total evidence at the later time will come from $\mathcal{E} = \{E_1, \ldots, E_m\}$, then

rationality requires that, if R is your updating plan, then R is a Bayesian plan for C—that is, for each E_i in $\mathcal{E}$, if $C(E_i) > 0$, then $R_i(-) = C(-|E_i)$.

To establish this, at least in the case in which $\mathcal{E}$ is a partition, the epistemic utility theorist appeals to a result by Hilary Greaves and David Wallace, building on work by Graham Oddie (Oddie, 1997; Greaves & Wallace, 2006). Here, for any possible world w, we write E_{i_w} for the member of the partition $\mathcal{E}$ to which w belongs. So, given an updating rule R, R_{i_w} is the credence function it requires you to adopt at world w, since you'll learn evidence E_{i_w} at that world. So we might naturally take $\mathfrak{EU}(R_{i_w}, w)$ to measure the epistemic utility of updating plan R at world w—it is the epistemic utility of the credence function this plan requires you to adopt if you learn the evidence that you'll learn at world w.

Theorem 3 *Suppose $\mathfrak{EU}$ is a strictly proper epistemic utility measure, C is a probabilistic credence function, and $\mathcal{E}$ is a partition.*

(I) *If R, R' are both Bayesian plans for C, then*

$$\sum_{w \in \mathcal{W}} C(w)\mathfrak{EU}(R_{i_w}, w) = \sum_{w \in \mathcal{W}} C(w)\mathfrak{EU}(R'_{i_w}, w)$$

(II) *If R is a Bayesian plan for C and R' is not, then*

$$\sum_{w \in \mathcal{W}} C(w)\mathfrak{EU}(R'_{i_w}, w) < \sum_{w \in \mathcal{W}} C(w)\mathfrak{EU}(R_{i_w}, w)$$

So, your credence function at the earlier time expects Bayesian updating rules to maximise expected epistemic utility: it assigns each Bayesian rule the same expected epistemic utility, and it assigns any other rule a lower expected epistemic utility. From this, we conclude the narrow scope version of Plan Conditionalization. It is narrow scope because it has the following form: If your credence function is ..., then rationality requires that your updating plan is .... We now meet a wide scope version of the norm, which has the following form: Rationality requires that, if you credence function is ..., then your updating plan is ....

Plan Conditionalization (wide scope) If you know that your total evidence at the later time will come from $\mathcal{E} = \{E_1, \ldots, E_m\}$, then rationality requires that, if C is your credence function at the earlier time, and if R is your updating plan, then R is a Bayesian plan for C—that is, for each E_i in $\mathcal{E}$, if $C(E_i) > 0$, then $R_i(-) = C(-|E_i)$.

To establish this, we appeal to the following result (Briggs & Pettigrew, 2020):

Theorem 4 *Suppose $\mathfrak{EU}$ is a strictly proper epistemic utility measure and $\mathcal{E}$ is a partition.*

(I) *If R is not a Bayesian plan for C, then there is an alternative credence function $C^\star$ and updating plan $R^\star$ such that $R^\star$ is a Bayesian plan for C, and for all w in $\mathcal{W}$,*

$$\mathfrak{EU}(C, w) + \mathfrak{EU}(R_{i_w}, w) < \mathfrak{EU}(C^\star, w) + \mathfrak{EU}(R^\star_{i_w}, w)$$

(II) *If R is a Bayesian plan for C, there is no alternative credence function $C^\star$ and alternative plan $R^\star$ such that, for all w in $\mathcal{W}$,*

$$\mathfrak{EU}(C, w) + \mathfrak{EU}(R_{i_w}, w) \leq \mathfrak{EU}(C^\star, w) + \mathfrak{EU}(R^\star_{i_w}, w)$$

So, if you plan to update in some way other than using a Bayesian updating rule, there is an alternative credence function and an alternative updating plan that, taken together, is guaranteed to be better than your credence function and updating plan, taken together. From this, together with Undominated Dominance, we conclude a wide scope version of Plan Conditionalization.

In brief...

In Section 3.1, we met Thomas Kelly's argument for permissivism about rational belief (Kelly, 2014). It is inspired by William James' permissivism about attitudes to epistemic risk (James, 1897). James held that we have two goals in our epistemic life: to believe truths, and not to believe falsehoods. How we weigh these against each other is determined by our attitudes to epistemic risk. If you are risk-averse, you place greater weight on avoiding getting things wrong, while if you are risk-inclined you place greater weight on getting things right. James held that many different ways of weighing these two goals are permissible.

We interpreted Kelly's argument using an account of epistemic value that has been explored by Kenny Easwaran (2016) and Kevin Dorst (2019). According to this, we encode these Jamesian attitudes to epistemic risk in our epistemic utilities—that is, in our measures of epistemic value. For the risk-averse, the disvalue of believing falsely is greater than the value of believing truly; for the risk-inclined, it is reversed; for the risk-neutral, they are the same. Kelly posits epistemic probabilities: that is, for each body of evidence and each proposition, he posits a unique objective, agent-independent probability that measures how likely that body of evidence makes that proposition. Combining this with Easwaran and Dorst's account of epistemic utility, we can then proceed as follows. For any individual, any total body of evidence, and any proposition about which they might have an option, we calculate the expected epistemic utilities of (i) believing the proposition, (ii) disbelieving the proposition, and (iii) withholding judgment about it, all from the point of view of the evidential probability of the proposition relative to the individual's total evidence. Kelly holds that rationality requires us to do whatever maximises expected epistemic utility from that point of view. And he notes that, since rationality permits a number of different attitudes to epistemic risk, which we then encode in our epistemic utilities, the same evidential probability might demand one person to believe a proposition and another to withhold judgment—the more risk-averse you are, the stronger your evidence will have to be to demand belief. From this, he concludes in favour of permissivism about rational belief.

In Section 3.2, we noted Sophie Horowitz's observation that a similar argument will not deliver permissivism in the credal case (Horowitz, 2017). This is because, in that case, it is generally agreed that all legitimate measures of epistemic utility must have a particular property—in the jargon, they must be *strictly proper*. But, for any strictly proper epistemic utility function and any evidential probability, the evidential probability expects itself to be epistemically better than any other credence. So it renders itself the unique rational response to the evidence, which gives impermissivism. So we concluded that, if we are to secure permissivism for rational credences using epistemic utility theory, we can't adopt Kelly's approach.

In the remainder of Section 3.2 and in Section 3.3, we considered these strictly proper epistemic utility functions in greater detail, and we met some of the central results of epistemic utility theory, which furnish arguments for various Bayesian norms for credences: the norm that says that your credences should obey the probability calculus (Probabilism); the norm that says that they should defer to the objective chances (Principal Principle); and the norm that says that you should plan to update your credences using Bayes' Rule (Plan Conditionalization).

4 Epistemic risk and picking priors I: the decision rule

In the previous chapter, we met Thomas Kelly's argument for epistemic permissivism for full beliefs. In that, he argued that permissivism about our attitudes to epistemic risk, encoded in our epistemic utilities, entails permissivism about the epistemic rationality of full beliefs. Unfortunately, as Horowitz points out, a similar approach will not work for credences, which are our topic here.

But all is not lost. After all, as has long been recognised in approaches to practical decision-making, and as I'll explain in detail in Section 4.1, there are two ways in which we can encode our attitudes to risk: we can encode them in our utilities, as Kelly would have us do; or we can encode them in the decision rules that we use when we face a decision problem. As we'll see below, economists and decision theorists first realised this when Maurice Allais presented a set of apparently rational risk-averse preferences that cannot be rationalised by encoding the attitudes to risk in the utilities while retaining the standard risk-neutral expected utility rule for decision-making.

In this chapter, I'll consider various candidate decision rules that we might use to pick our prior credences—that is, the credences we have at the beginning of our epistemic life, when we have no evidence and no other credences. I'll explain how we might distinguish these different decision rules by appealing to axioms that characterise the preference orderings that give rise to them. As I discuss those axioms, I'll argue in favour of some and against others. The ones that I favour characterise a new family of decision rules called the *generalised Hurwicz criterion*. When we determine which prior credences it is rational for an individual to have, it is one of these decision rules, along with our epistemic utilities for credences, that we should apply.

Thus, our approach, like Kelly's, belongs to epistemic utility theory, where we take the rational credences to be those deemed rational by the correct decision rule when the options in question are our doxastic states, and the utility of each option is its epistemic utility. And, like Kelly, the priors that are rationally permissible for a particular individual are determined by their attitudes to risk. However, unlike Kelly, those attitudes to risk will appear not in our epistemic utilities, but in our decision rule. In the following chapter, I'll derive the consequences of applying these rules to our doxastic decision problem.

4.1 Risk-sensitive decision-making under risk

In this section, I'd like to explain how we might incorporate risk into utilities in practical sphere, and then explain why economists and philosophers historically thought it necessary to go beyond this and introduce risk-sensitive decision rules. I will be concerned with the case of *decision-making under risk*, which covers those decision problems for which we know the objective probabilities of the different possible states of the world on which the available options are defined. This is not our situation when we pick our priors. At that point, we know nothing of the objective probabilities. Picking priors is a case of *decision-making under uncertainty*, which covers those decision problems for which we do not know the objective probabilities. I turn to decision-making under uncertainty in Section 4.2; if you wish, you can go there now without loss.

We'll begin with the simplest possible example. I am about to flip a fair coin. I make you an offer: pay me £30, and I will pay you £100 if the coin lands heads and nothing if it lands tails—that is, if you accept my offer and the coin lands heads, you will make a net gain of £70, while if you accept and it lands tails, you'll lose £30; if you don't accept, you'll lose nothing and gain nothing either way. You choose to reject my offer. But why? After all, the expected monetary payout of accepting my offer is £ $\left((\frac{1}{2} \times 70) + (\frac{1}{2} \times -30)\right) =$ £20, and this is greater than the expected monetary payout of rejecting it, which is £0. Have you been irrationally risk-averse? Have you placed more weight on the badness of the possible £30 loss than is warranted, and less weight on the goodness of the possible £70 gain?

There are two ways to rationalise your decision. On the first, we retain the standard decision rule that demands that you maximise your expected utility; we encode your attitude to risk in your utility function; and we show that you in fact maximise expected utility by rejecting my offer, even though you don't maximise expected monetary payout. On the second, we abandon the expected utility rule; we introduce a new decision rule to govern these situations; and we show that rejecting my offer is exactly what that rule demands, or at least permits.

4.1.1 Expected utility theory and the diminishing marginal utility of money

Let's present the first approach in a little more detail. The key point is this: we noted that the expected monetary payout of accepting my offer exceeds the expected monetary payout of rejecting it; but our decision-principle says nothing about expected monetary payouts, and instead talks only of expected utility. And, as Daniel Bernoulli and Gabriel Cramer noted in the very earliest discussions of expected utility theory, I can rationalise your decision to reject my offer by ascribing to you a certain sort of utility function (Bernoulli, 1738 [1954]). In particular, we need a utility function u such that

$$\frac{1}{2}u(\text{Gain £70}) + \frac{1}{2}u(\text{Lose £30}) < u(\text{Status quo}).$$

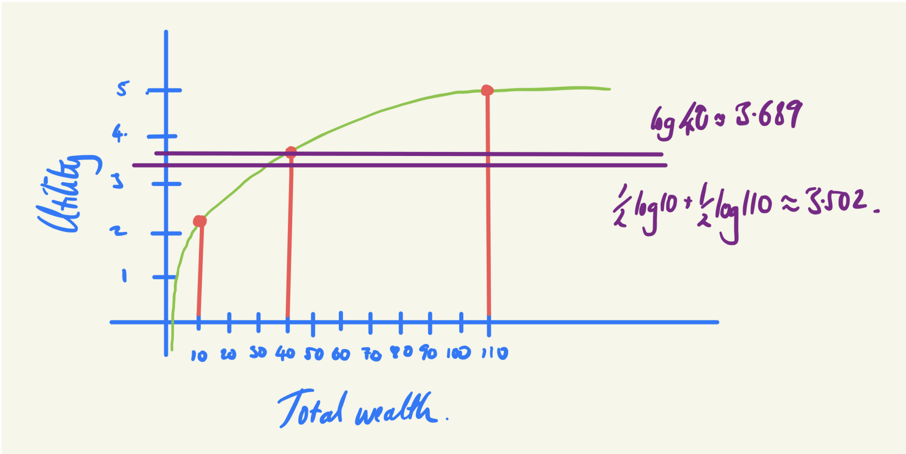

Figure 4: Utility as a logarithmic function of wealth

And, as Bernoulli and Cramer observed, a utility function that is a sufficiently concave function of your total wealth as measured in monetary units will do the job. For instance, to borrow Bernoulli's own suggestion, perhaps your utility for an outcome in which your total wealth is £n is $\log n$. And perhaps your current wealth is £40.

If that's the case, then your total wealth if you were to accept my offer and the coin were to land heads would be £110, while your total wealth if you were to accept and the coin were to lands tails would be £10. Then your expected utility for accepting my offer is $\frac{1}{2} \log 110 + \frac{1}{2} \log 10 \approx 3.502$ while your expected utility for rejecting it is $\log 40 \approx 3.689$. So, you maximise your expected utility by rejecting my offer. What's more, it seems reasonable for you to value money in the way that Bernoulli's logarithmic utility function suggests. After all, money is what economists call a *dependent good*—how valuable a certain amount of it is depends on how much you have already. And, in particular, it has diminishing marginal utility. That is, each extra unit of money adds less utility than the previous one. The more money you have, the less extra utility you gain by receiving a certain amount more. So giving £10 to a millionaire increases their utility less than giving £10 to someone living in poverty increases theirs. And that's precisely what the logarithmic utility function suggests. If my total wealth is £100 and I receive £10, then my utility increases from $\log 100 \approx 4.61$ to $\log 110 \approx 4.7$—an increase of around 0.09. Whereas if my total wealth is £1,000,000 and I receive £10, then my utility increases from $\log 1,000,000 \approx 13.81551$ to $\log 1,000,010 \approx 13.81552$—an increase of around 0.00001.

4.1.2 Risk-weighted expected utility and the Allais preferences

On the approach just sketched, we take a risk-neutral decision rule—the one that says you should maximise your expected utility—and combine it with a risk-averse utility function—the logarithmic one, or some other sufficiently concave function—and thereby rationalise your decision to reject my offer. But we might rationalise that decision in the opposite way. We might abandon the risk-neutral decision rule in favour of a risk-averse one, and ascribe to you a utility function that is linear in money. What would a risk-averse decision rule look like? As we'll see below, they come in two varieties: those that apply when probabilities, either subjective or objective or evidential, are available for the outcomes of the decision; and those that apply when no such probabilities are available. In the epistemic case, we'll be more interested in the latter variety. But for the moment, let's meet one of the former variety, due to John Quiggin and Lara Buchak (Quiggin, 1982, 1993; Buchak, 2013).

We say the expected utility rule is risk-neutral because the weight it assigns to each outcome is just the probability of that outcome. Buchak's risk-weighted expected utility rule allows us to use other weights for the outcomes—in particular, it allows us to give greater weight to worse-case scenarios and less weight to best-case scenarios than expected utility theory allows. To see how it works, consider the following alternative way of describing the expected utility of the offer that I make to you. As we usually present expected utility, we take each of the two outcomes—the first where the coin lands heads and you make a net gain of £70, the second where the coin lands tails and you make a net loss of £30—we weight them by the probabilities of these outcomes, and we add these weighted utilities together to give the expected utility. That is:

$$\text{Expected utility of accepting the offer} = \frac{1}{2}u(\text{Gain £70}) + \frac{1}{2}u(\text{Lose £30})$$

But we might equally describe it as follows: we start with the worst-case outcome—you make a net loss of £30—and weight its utility by the probability you'll obtain at least that much utility, which is obviously 1; then we take the amount of extra utility over and above that worst-case utility you'd get from the second-worst case outcome—where you make a net gain of £70—and weight that extra utility by the probability you'll get that as well, which is $\frac{1}{2}$ in this case; then you add them up. In other words:

$$\text{Expected utility of accepting the offer} = u(\text{Lose £30}) + \frac{1}{2}(u(\text{Gain £70}) - u(\text{Lose £30}))$$

Now, according to the risk-weighted expected utility rule, we do something similar to the second calculation, but instead of weighting the extra utility you'll gain from the second-worst outcome by the probability you'll get that extra utility, you weight it by some transformation of that probability. This transformation is effected by your *risk function* r, which takes the probabilities you'd use in the expected utility rule and returns the weightings you'll use in the risk-weighted expected utility rule. In particular:

$$\text{Risk-weighted expected utility of accepting the offer} =$$
$$u(\text{Lose £30}) + r(\frac{1}{2})(u(\text{Gain £70}) - u(\text{Lose £30}))$$

We insist that:

(i) $r(0) = 0$ and $r(1) = 1$,

(ii) if $0 \leq p \leq 1$, then $0 \leq r(p) \leq 1$, and

(iii) r is a strictly increasingly function, so that if $0 \leq p < q \leq 1$, then $0 \leq r(p) < r(q) \leq 1$.

But otherwise we have a lot of freedom to define r. If $r(p) < p$ for all $0 < p < 1$, then we say that r is risk-averse throughout its range. That's because it gives less weight to the better case scenario than expected utility theory does—in our case, it gives $r(\frac{1}{2}) < \frac{1}{2}$. If $r(p) > p$ for all $0 < p < 1$, then we say that r is risk-inclined throughout its range. That's because it gives more weight to the better case scenario than expected utility theory does—in our case, it gives $r(\frac{1}{2}) > \frac{1}{2}$. And of course, if $r(p) = p$ for all $0 < p < 1$, then risk-weighted expected utility is just expected utility.

Now, suppose your utility is linear in money. So your utility for an outcome in which your total wealth is £n is n. And suppose your current wealth is £40. Then the risk-weighted expected utility of accepting the offer is:

$$\text{Risk-weighted expected utility of accepting the offer} =$$
$$u(\text{Lose £30}) + r(\frac{1}{2})(u(\text{Gain £70}) - u(\text{Lose £30})) =$$
$$10 + r(\frac{1}{2})100$$

And the risk-weighted expected utility of declining the offer is just 40, since choosing that will guarantee your current wealth level of £40, and so the worst-case scenario and the second-worst-case scenario are the same and you receive £40 in each. So risk-weighted expected utility will demand that you decline my offer if $10 + r(\frac{1}{2})100 < 40$, which will hold iff $r(\frac{1}{2}) < \frac{3}{10}$. So, for instance, if your risk function is $r(p) = p^2$, then $r(\frac{1}{2}) = \frac{1}{4} < \frac{3}{10}$, and you should decline my offer.

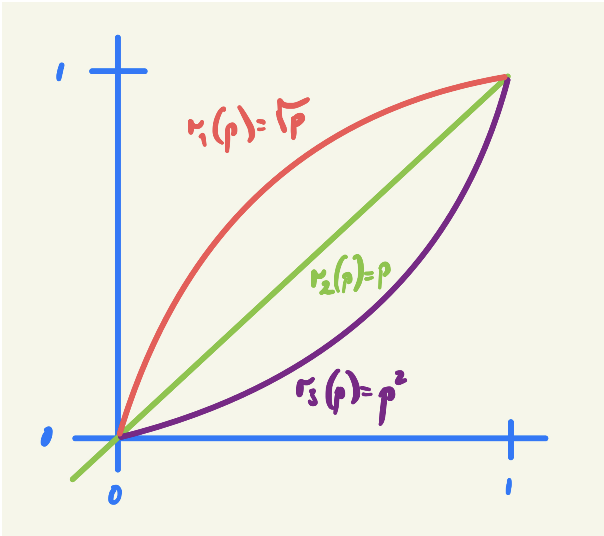

Figure 5: Three risk functions: r_1 is risk-inclined, r_2 is risk-neutral, and r_3 is risk-averse.

Upon first meeting the risk-weighted expected utility rule, you might wonder why we go to the bother of creating a whole new decision rule, when we have seen that we can instead encode our attitudes to risk in our utility function. The reason is that there are certain combinations of choices we might make that seem perfectly reasonable, but which cannot be rationalised using Bernoulli's trick of picking an appropriate utility function. The classic example was first described by Maurice Allais (1953). He described four lotteries, a, b, c, and d. Each has a million tickets. But the payouts from the tickets are different in the four lotteries. He then asked which you would choose if you were offered a random ticket from a or a random ticket from b; and which you would choose if you were offered a random ticket from c or a random ticket from d. He noted that it seems reasonable to choose a over b, and d over c. But he also showed that there is no utility function you can ascribe to an individual such that a has higher expected utility than b and d has higher expected utility than c relative to it. That is, for any utility function, if a has higher expected utility than b, then c has higher expected utility than d. Here are the four lotteries:

Tickets	a	b	c	d
1-89	£1m	£1m	£0m	£0m
90	£1m	£0m	£1m	£0m
91-100	£1m	£5m	£1m	£5m

Now, notice that a and b give the same outcome for the first eighty-nine tickets; and so do c and d. So, according to expected utility theory, which option you prefer out of a and b depends only on your attitudes to their different outcomes for the other eleven tickets; and similarly for c and d. But, c has exactly the same outcomes for those remaining eleven tickets as a, and d has exactly the same outcomes as b. So your expected utility for a is greater than your expected utility for b iff your expected utility for c is greater than your expected utility for d.[15] The upshot: expected utility theory cannot rationalise a preference for a over b and d over c.

For this reason, we must introduce risk-weighted expected utility, or something similar. After all, if you choose a over b but d over c, this suggests that you are risk-averse. The worst-case scenario of b—namely, £0m—is worse than the worst-case scenario of a—namely, £1m. But this is not true of d and c, which have the same worst-case scenario—namely, £0m. And we can see that, even with a utility function that is linear in monetary wealth, and even for someone with no initial wealth, risk-weighted expected utility can rationalise this behaviour. For instance, let $r(p) = p^3$. Then here are the risk-weighted utilities of the four lotteries:

- a: $1 \times 1,000,000 = 1,000,000$
- b:

$$(1 \times 0) + (r(0.99) \times (1,000,000 - 0)) + \\ (r(0.1) \times (5,000,000 - 1,000,000)) = 974,299$$

- c: $(1 \times 0) + (r(0.11) \times (1,000,000 - 0)) = 1,331$
- d: $(1 \times 0) + (r(0.1) \times (5,000,000 - 0)) = 5,000$

The risk-weighted expected utility of a exceeds that of b, and the risk-weighted utility of d exceeds that of c, just as we wanted.

[15] In symbols, let u_0, u_1, u_5 be your utilities for the outcomes £0m, £1m, £5m, respectively. Then

$$\begin{array}{rcl} \text{Expected utility of } a & > & \text{Expected utility of } b \\ & \text{iff} & \\ 0.89u_1 + 0.01u_1 + 0.1u_1 & > & 0.89u_1 + 0.01u_0 + 0.1u_5 \\ & \text{iff} & \\ 0.01u_1 + 0.1u_1 & > & 0.01u_0 + 0.1u_5 \\ & \text{iff} & \\ 0.89u_0 + 0.01u_1 + 0.1u_1 & > & 0.89u_0 + 0.01u_0 + 0.1u_5 \\ & \text{iff} & \\ \text{Expected utility of } c & > & \text{Expected utility of } d \end{array}$$

4.2 Risk-sensitive decision-making under uncertainty

So much for risk-sensitive decision rules that cover those cases when objective probabilities are available. While Allais' preferences showed the necessity of developing such rules for those cases, risk-sensitive decision rules had been developed a few years earlier for those cases in which the objective probabilities are not available. They came during a flurry of activity at the Cowles Commission in Chicago at the end of the 1940s, and the work was undertaken by a group of young researchers who would go on to make great contributions in their respective fields: the mathematicians Abraham Wald and John Milnor, and the economists Leonid Hurwicz and Kenneth Arrow. From what I can tell, this group didn't set out to investigate risk-sensitive decision-making specifically. Rather, they were interested in how we might choose in the absence of any evidence, and specifically in the absence of any knowledge of the objective chances.

There are two species of decision rule that govern decision-making under uncertainty: those that begin by setting endogenous probabilities—in particular, credences—and then use those to choose between options; and those that do not use probabilities at all. We'll consider four of the first species and three of the second. In the first group:

(i) *subjective Bayesianism*, which says that you may pick any probability distribution and then maximise expected utility with respect to that;

(ii) *risk-weighted subjective Bayesianism*, which says that you may pick any probability distribution and then maximise risk-weighted expected utility with respect to that;

(iii) *objective Bayesianism*, which says that you should maximise expected utility with respect to the uniform distribution;

(iv) *risk-weighted objective Bayesianism*, which says that you should maximise risk-weighted expected utility with respect to the uniform distribution.

And in the second group, which does not set endogenous probabilities, we'll consider:

(v) *the Maximin rule*, which says that you should pick the option whose worst-case utility is highest;

(vi) *the Hurwicz criterion*, which says you should set a weighting for the worst outcome and a weighting for the best outcome, and then you should pick the option where the weighted average of the worst utility and best utility is highest;

(vii) *the generalised Hurwicz Criterion*, which says you should set weightings for the worst, second-worst, ..., second-best, and best outcomes, and then you should pick the option where the weighted average of the worst, second-worst, ..., second-best, and best utilities is highest.

We are interested in these decision rules because they are designed to cover exactly the sort of decision problem that epistemic utility theory takes us to face when we set our prior credences. In that case, as in the cases these are designed to cover, we have no evidence; and, *a fortiori*, we have no access to the objective probabilities that govern the possible worlds over which our options are defined.

Here's the framework. The subject of our study is an individual facing a decision problem. A decision problem consists of a series of available options between which the individual must choose. The options are specified at a particular level of grain, which is determined by the set of possible worlds that the individual uses. Each option is determined by the utility that it assigns to each possible world in the individual's set. A decision rule takes the available options in a decision problem and divides them into those that rationality permits and those that it does not permit. Each of the decision rules we'll describe works in the same way: it begins by assigning a numerical score to each option; these numerical scores determine the individual's preference ordering over the options, with one option preferred to another just in case the first one's score is higher than the second one's; and then the decision rule says that an option in a decision problem is irrational if there is another available option that is strictly preferred.

- *Possible worlds* Let $\mathcal{W}$ be the set of individual's set of possible worlds. At a minimum, these must be specified in enough detail that they fix all the aspects of the outcomes of a choice that determine its utility for our agent. For instance, if the agent is a pure hedonist egoist and values only their own pleasure, each w in $\mathcal{W}$ would have to specify exactly the amount of pleasure that they enjoy at that world. Or, in the case of the doxastic decision problem that interests us, each w in $\mathcal{W}$ would have to specify everything that determines the epistemic utility of each credence function. Throughout, we'll assume that the set of possible worlds is finite. In particular, we'll let $\mathcal{W} = \{w_1, \ldots, w_n\}$.

- *Options* Given a set of possible worlds $\mathcal{W}$, an option is represented by a function from $\mathcal{W}$ into the real numbers. If a is an option and $\mathcal{W}$ is a possible world in $\mathcal{W}$, then $a(w)$ is the utility of a at $\mathcal{W}$.[16]

[16]We fix a single numerical scale on which the utility of all options are measured. As Hurwicz himself notes, we can use the von Neumann-Morgenstern representation theorem to determine the numerical utilities that our individual assigns to the outcome of an option at a world (von Neumann & Morgenstern, 1947). That is, we can appeal to our preferences when we know the objective probabilities of the various outcomes to discover our utilities

- *Preferences* Our individual has a preference ordering over options. This is an ordering $\preceq$ on the set of options.

  The preference ordering $\preceq$ is what is sometimes called the weak or non-strict preference ordering. This distinguishes it from the strong or strict preference ordering $\prec$. Given $\preceq$, we write $a \prec b$ if $a \preceq b$ but $b \npreceq a$, and we write $a \sim b$ if $a \preceq b$ and $b \preceq a$. We say that a and b are comparable if $a \preceq b$ or $b \preceq a$ or both.

We are now ready to meet our decision rules. In fact, in many cases, we meet a family of decision rules. Each member of the family corresponds to a different choice of some further ingredient—sometimes a probability function over the possible worlds; sometimes some other way of weighting the utilities of the options at different worlds.

4.2.1 Subjective Bayesianism

Given an option a and a probability assignment $0 \leq p_1, \ldots, p_n \leq 1$ with $\sum_{i=1}^{n} p_i = 1$, which we denote P, define the expected utility of a relative to P as follows:

$$\text{Exp}^P(a) := p_1 a(w_1) + \ldots + p_n a(w_n)$$

Subjective Bayesianism says that you should order options by their expected utility relative to a probability assignment of your choice. Thus, given P,

$$a \preceq^P_{sb} a' \Leftrightarrow \text{Exp}^P(a) \leq \text{Exp}^P(a')$$

And the corresponding decision rule says that you should pick a probability assignment P and then, having done that, it is irrational for you to choose a if there is a' such that $a \prec^P_{sb} a'$.

4.2.2 Objective Bayesianism

Let $P^\dagger$ be the uniform distribution. That is,

$$P^\dagger(w_1) = \ldots = P^\dagger(w_n) = \frac{1}{n}$$

Objective Bayesianism says that you should order options by their expected utility relative to the uniform distribution. Thus,

$$a \preceq_{ob} a' \Leftrightarrow \text{Exp}^{P^\dagger}(a) \leq \text{Exp}^{P^\dagger}(a')$$

in the outcomes, and then use those when we are choosing in the absence of information about the objective probabilities.

In our representation theorems below, we'll assume a rich set of options—every function from the worlds to the reals will count as an option. However, we need not claim that they are all available to the agent. Rather, the agent's preferences are defined over all possibilities, so that when a subset of these becomes available to them in a decision problem that they face, they can appeal to their preference ordering to make their choice.

And the corresponding decision rule says that it is irrational to choose a if there is a' such that $a \prec_{ob} a'$. So we have our decision rule: a is irrational if there is a' such that $\mathrm{Exp}^{P^\dagger}(a) < \mathrm{Exp}^{P^\dagger}(a')$.

4.2.3 Risk-weighted subjective Bayesianism

Given an option a, a probability assignment $0 \leq p_1, \ldots, p_n \leq 1$ with $\sum_{i=1}^n p_i = 1$, and a risk function $r : [0,1] \to [0,1]$, which is strictly increasing and continuous with $r(0) = 0$ and $r(1) = 1$, define the risk-weighted expected utility of a relative to P and r as follows: if

$$a(w_{i_1}) \leq a(w_{i_2}) \leq \ldots \leq a(w_{i_n})$$

then

$$\begin{aligned}\mathrm{RExp}^{P,r}(a) := \\ & a(w_{i_1}) + r(p_{i_2} + \ldots + p_{i_n})(a(w_{i_2}) - a(w_{i_1})) + \\ & r(p_{i_3} + \ldots + p_{i_n})(a(w_{i_3}) - a(w_{i_2})) + \ldots + \\ & r(p_{i_{n-1}} + p_{i_n})(a(w_{i_{n-1}} - w_{i_{n-2}})) + \\ & r(p_{i_n})(a(w_{i_n}) - a(w_{i_{n-1}}))\end{aligned}$$

Risk-Weighted Subjective Bayesianism says that you order options by their risk-weighted expected utility relative to a probability assignment and risk function of your choice (Buchak, 2013). Thus, given P and r:

$$a \preceq_{rs}^{P,r} a' \Leftrightarrow \mathrm{RExp}^{P,r}(a) \leq \mathrm{RExp}^{P,r}(a')$$

And the corresponding decision rule says that you should pick a probability assignment P and a risk function r and then, having done that, it is irrational to choose a if there is a' such that $a \prec_{rs}^{P,r} a'$.

4.2.4 Risk-Weighted Objective Bayesianism

Risk-Weighted Objective Bayesianism says that you should order options by their risk-weighted expected utility relative to the uniform probability assignment. Your preference ordering should be such that there is a risk function r for which:

$$a \preceq_{ro}^{r} a' \Leftrightarrow a \preceq_{rs}^{P^\dagger,r} a' \Leftrightarrow \mathrm{RExp}^{P^\dagger,r}(a) \leq \mathrm{RExp}^{P^\dagger,r}(a')$$

That is, you should order options by their risk-weighted expected utility score relative to the uniform distribution and a risk function of your choice.

And the corresponding decision rule says that you should pick a risk function and that, having done that, it is irrational to choose a if there is a'

such that $a \prec a'$. So we have our decision rule: for someone with a risk function r, a is irrational if there is a' such that $\text{RExp}^{P^\dagger,r}(a) \leq \text{RExp}^{P^\dagger,r}(a')$.

This completes our presentation of the four accounts of preference orderings that take probabilities as one of their inputs. Next, we turn to those accounts that do not.

4.2.5 Wald's Maximin rule

Given an option a, define the *Wald score of a* as follows:

$$W(a) := \min_{w \in \mathcal{W}} a(w)$$

That is, $\mathcal{W}(a)$ is the minimum utility that a receives; the utility it receives at the world at which it receives lowest utility (Wald, 1945).

Maximin says that you should order options by their Wald score. Thus,

$$a \preceq_{mm} a' \Leftrightarrow W(a) \leq W(a')$$

And the corresponding decision rule says that it is irrational to choose a if there is a' such that $a \prec_{mm} a'$.

The problem with Wald's rule is that it demands an excess of caution. It puts all of its weight on the worst case and thereby ignores the second-worst case, the third-worst, etc., all the way up to the best case. For instance, I might offer you two options, a and b. Their outcome depends on which of two worlds we're in: w_1 and w_2. You know nothing of the objective chance of these two worlds; you know nothing at all that bears on which is actual. Here are the outcomes in utiles, the unit of utility, where ε is a very very small finite quantity:

	a_1	b_1
w_1	1	$1 - \varepsilon$
w_2	$1 + \varepsilon$	$1,000$

Wald's Maximin requires you to choose a_1, even though b_1 looks at the very least permissible.

Another problem is that, on its own, Maximin permits you to choose a weakly dominated option. Consider, for instance a_2 and b_2:

	a_2	b_2
w_1	1	1
w_2	$1 + \varepsilon$	$1,000$

The options a and a' have the same Wald score—namely, 1. So Wald's rule demands that our agent is indifferent between them—that is, $a_2 \sim b_2$. And so, in a decision between just those options, it permits both. But b_2 is at least as good as a_2 at all worlds, and better at some—that is, in the jargon of decision theory, b_2 weakly dominates a_2. And surely it is irrational to choose a weakly dominated option.

Now, we can save Wald's decision rule from this objection by saying that it merely gives a sufficient condition for irrationality: it says only that an option is irrational if there is an alternative with a lower Wald score. It does not say that this is the only route to irrationality. So we might combine Wald's decision rule with a rule that says that being weakly dominated also suffices for irrationality. Thus, a is irrational if *either* (i) there is a' with $a \preceq_{mm} a'$ *or* (ii) there is a' that weakly dominates a *or* both (i) and (ii). But this does not fix the first problem. Whether or not it is permissible to be as risk-averse as Wald's rule requires you to be, it is certainly not mandatory.

4.2.6 Hurwicz's Criterion of Realism

Noting that Wald's decision rule was too cautious and unreasonably ignored outcomes other than the very worst, Hurwicz introduced an alternative that pays attention not only to the worst-case outcome but also to the best-case outcome (Hurwicz, 1951, 1952). It asks you to score an option not by its worst-case utility alone, but by a weighted average of its worst-case and the best-case utilities. That weighting, Hurwicz thought, measured how optimistic or pessimistic you are—hence the name, 'the criterion of realism'. The optimist gives more weight to the best-case utility; the pessimistic gives more to the worst-case. I think it's more reasonable to see this not as a reflection of pessimism or optimism, which seems to make it an attitude about how the world is, but as a reflection of risk-sensitivity. The more risk-averse you are, the more weight you'll give to the worst-case scenario; the more risk-inclined, the more weight you bestow upon the best-case scenario.

Given an option a and a weight $0 \leq \lambda \leq 1$, define the *Hurwicz score of a relative to* λ as follows:

$$H^{\lambda}(a) := \lambda \max_{w \in \mathcal{W}} a(w) + (1 - \lambda) \min_{w \in \mathcal{W}} a(w)$$

That is, $H^{\lambda}(a)$ is the weighted average of the minimum utility a receives and the maximum utility it receives.

The Hurwicz Criterion says that you should order options by their Hurwicz score relative to a weighting λ of your choice. Thus, given λ,

$$a \preceq^{\lambda}_{hc} a' \Leftrightarrow H^{\lambda}(a) \leq H^{\lambda}(a')$$

And the corresponding decision rule says that you should pick your Hurwicz weight λ and then, having done that, it is irrational to choose a if there is a' such that $a \prec^{\lambda}_{hc} a'$.

The idea is that you specify your attitude to risk by specifying your Hurwicz weight λ, and then the decision rule specifies what options are irrational for you. The higher λ is, the more risk-inclined; the lower it is, the more risk-averse.

First, let's note that, for a huge range of weights, the Hurwicz Criterion delivers the verdict we want in the cases above that Maximin got wrong. Providing $\lambda > \frac{\varepsilon}{999}$, then

$$H^\lambda(a_1) = (1+\varepsilon)\lambda + (1-\lambda) < 1,000\lambda + (1-\varepsilon)(1-\lambda) = H^\lambda(b_1)$$

So, unless you're extremely risk-averse, b_1 is the rational choice, which matches our intuitive verdict. Similarly, providing $\lambda > 0$, then

$$H^\lambda(a_2) = (1+\varepsilon)\lambda + (1-\lambda) < 1,000\lambda + (1-\lambda) = H^\lambda(b_2)$$

So, unless you are maximally risk-averse, b_2 is the rational choice. Note, of course, that this is because, when $\lambda = 0$, the Hurwicz score and the Wald score coincide: $H^0(a) = W(a)$, for any option a.

However, analogous problems haunt Hurwicz's Criterion. Consider the following choice:

	a_3	b_3
w_1	1	$1-\varepsilon$
w_2	$1+\varepsilon$	$1,000-\varepsilon$
w_3	$1,000+\varepsilon$	$1,000$

Then, for any weighting, the Hurwicz score of a_3 is higher than the Hurwicz score of b_3. And yet, as before, it seems that b_3 is at least rationally permissible. Also, consider the following choice:

	a_4	b_4
w_1	1	1
w_2	$1+\varepsilon$	$1,000-\varepsilon$
w_3	$1,000$	$1,000$

Then b_4 weakly dominates a_4, but their Hurwicz scores are the same, and so in a decision between only the two of them, both are rationally permitted. Though note that, again, we can simply add to the decision rule that weakly dominated options are irrational.

4.2.7 The generalised Hurwicz Criterion

Now, we noted above that the Hurwicz Criterion is essentially a weakening of the Maximin rule: that is, Maximin is a special case of the Hurwicz Criterion, and so the latter permits anything that the former permits. The natural reaction to the problems that we've identified with both rules is to move further in the same direction. This leads us to my generalisation of the Hurwicz Criterion.

Suppose $\mathcal{W} = \{w_1, \ldots, w_n\}$ is the set of possible worlds over which our options are defined. Given an option a and a sequence of weights $0 \leq \lambda_1, \ldots, \lambda_n \leq 1$ with $\sum_{i=1}^{n} \lambda_i = 1$, which we denote Λ, define the *generalised Hurwicz score of a relative to* Λ as follows: if

$$a(w_{i_1}) \geq a(w_{i_2}) \geq \ldots \geq a(w_{i_n})$$

then

$$H^{\Lambda}(a) := \lambda_1 a(w_{i_1}) + \ldots + \lambda_n a(w_{i_n})$$

That is, $H^{\Lambda}(a)$ is the weighted average of all the possible utilities that a receives, where λ_1 weights the highest utility, λ_2 weights the second-highest, and so on.[17] The generalised Hurwicz Criterion says that you should order options by their generalised Hurwicz score relative to a sequence Λ of weightings of your choice. Thus, given Λ,

$$a \preceq^{\Lambda}_{ghc} a' \Leftrightarrow H^{\Lambda}(a) \leq H^{\Lambda}(a')$$

And the corresponding decision rule says that you should pick your Hurwicz weights Λ and then, having done that, it is irrational to choose a if there is a' such that $a \prec^{\Lambda}_{ghc} a'$.

A more risk-averse agent will pick a sequence of weights $\lambda_1, \lambda_2, \ldots, \lambda_n$ in which the early weights in the sequence, which weight the better cases, are less than the later weights, which weight the worse cases; and a risk-inclined agent will do the opposite. A risk-neutral agent will assign equal weight to all cases.

Notice that GHC permits reasonable verdicts for the examples given above: if $\lambda_2 > \frac{\varepsilon}{999-\varepsilon}$, then

$$H^{\Lambda}(a_3) = \lambda_3 + \lambda_2(1+\varepsilon) + \lambda_1(1000+\varepsilon) < \lambda_3(1-\varepsilon) + \lambda_2(1000-\varepsilon) + \lambda_1(1000) = H^{\Lambda}(b_3)$$

And, if $\lambda_3 < 1$, then

$$H^{\Lambda}(a_4) = \lambda_3 + \lambda_2(1+\varepsilon) + \lambda_1(1000) < \lambda_3 + \lambda_2(1000-\varepsilon) + \lambda_1(1000) = H^{\Lambda}(b_4)$$

[17] It is important to convince yourself that $H^{\Lambda}(a)$ is well-defined. After all, it's possible that

$$a(w_{i_1}) \geq a(w_{i_2}) \geq \ldots \geq a(w_{i_n})$$

and

$$a(w_{i_2}) \geq a(w_{i_1}) \geq \ldots \geq a(w_{i_n})$$

if $a(w_{i_1}) = a(w_{i_2})$. But then of course

$$\lambda_1 a(w_{i_1}) + \lambda_2 a(w_{i_2}) + \ldots + \lambda_n a(w_{i_n}) = \lambda_1 a(w_{i_2}) + \lambda_2 a(w_{i_1}) + \ldots + \lambda_n a(w_{i_n})$$

as required.

And indeed, providing all weights are positive—that is, $0 < \lambda_1, \ldots, \lambda_n < 1$—then if a' weakly dominates a, then $H^{\Lambda}(a) < H^{\Lambda}(a')$.

4.3 Characterizing our rules

In this section, we build on John Milnor's classic paper, 'Games Against Nature', to characterise the seven families of decision rule we have described (Milnor, 1954). That is, we will describe certain properties that a preference orderings might or might not have; and then, for each of the families of rules described, we'll find a set of those properties and show that a preference ordering belongs to the family in question if, and only if, it has all of the properties in this set. This will allow us to tell between the different families of rules by considering the properties that distinguish them.

While we will present combinations of axioms that characterise each of the families of decision rules presented above, our goal is to determine the decision rule that we should apply to discover which prior credences are rationally required. So, when we motivate certain axioms below, we motivate them as requirements on whichever decision rule governs that particular decision problem.

With this in mind, it is worth noting a theme that will run throughout: it is a certain sort of motivation for restricting certain axioms in a particular way. There are many accounts of when a family of decision rules permits risk-sensitivity. Here is mine: to permit risk-sensitivity, a family of decision rules must permit you to take the utility of an option a at a world $\mathcal{W}$ to contribute differently to the evaluation of that option depending on the position in the ordering of the worlds by utility obtained by a that $\mathcal{W}$ occupies. For instance, if $a(w_1) = 0$, $a(w_2) = 8$, and $b(w_1) = 24$, $b(w_2) = 8$, then both obtain the same utility at world w_2, but that is the best case for a and the worst case for b, and so a risk-sensitive decision theory should permit you to take the shared utility at w_2 to contribute differently to the overall evaluation of a and the overall evaluation of b. This will come up again and again in what follows. If we follow William James in permitting many different attitudes to epistemic risk, and we wish to encode those attitudes in our decision rule, rather than our epistemic utilities, we must ensure that our axioms don't rule out permissible attitudes to risk.

To give our characterizations, we need some notation:

- We will sometimes denote an option by the n-tuple of its utility values at the n different worlds. Thus, $a = (a(w_1), \ldots, a(w_n))$.
- Given a real number u, we write $\overline{u}$ for the constant option that has utility u at all worlds. That is, $\overline{u} = (u, \ldots, u)$.[18]

[18] As mentioned above, we assume the set of options over which your preference ordering is defined is very rich.

- If a, a' are options, then define $a + a'$ to be the option with $(a + a')(w) = a(w) + a'(w)$, for all w in $\mathcal{W}$. Thus,

$$a + a' := (a(w_1) + a'(w_1), \ldots, a(w_n) + a'(w_n))$$

- If a is an option and m, k are real numbers, then define $ma + k$ to be the option with $(ma + k)(w) = m \times a(w) + k$, for all w in $\mathcal{W}$. Thus,

$$ma + k := (ma(w_1) + k, \ldots, ma(w_n) + k)$$

- We say that *a and b order the worlds in the same way* if $a(w_i) \leq a(w_j)$ iff $b(w_i) \leq b(w_j)$ for all w_i, w_j in $\mathcal{W}$.

- We say that a and b are comparable if $a \preceq b$ or $b \preceq a$.

4.3.1 The axioms

We now present some axioms that govern preference orderings $\preceq$.

(A1) **Reflexivity** $\preceq$ is reflexive. That is, $a \preceq a$, for all a.

(A2) **Transitivity** $\preceq$ is transitive. That is, if $a \preceq b$ and $b \preceq c$, then $a \preceq c$, for all a, b, c.

These aren't controversial in this context. I won't say anything more about them.

(A3) **Weak Mixture Continuity** If $a \prec b \prec c$, and a, b, c order the worlds in the same way, then there is $0 < \alpha < 1$ such that $\alpha a + (1 - \alpha)c \sim b$.

(A3*) **Strong Mixture Continuity** If $a \prec b \prec c$, then there is $0 < \alpha < 1$ such that $\alpha a + (1 - \alpha)c \sim b$.

These requirements are more controversial. A natural way to justify them is assume to that I should set my preferences over the options by assigned to each option a numerical score and to order them by their scores. (As we noted above, this is an approach shared by all of the families of decision rules listed above—expected utilities, risk-weighted and otherwise, subjective or objective; Wald scores; and Hurwicz scores, generalised or not.) We then assume that this score is a continuous function of the options: that is, for any option a, and for any maximum difference ε between scores, there is some distance from a such that, providing an option b is within that distance from a, the score of b is within ε of the score of a. And so, in particular, the score of the mixtures $\alpha a + (1 - \alpha)c$ is a continuous function of α. But, since the score of a lies below b and the score of c lies above, the Intermediate Value Theorem tells us that there must be some $0 < \alpha < 1$ for which the score of $\alpha a + (1-\alpha)c$ is the same as the score of b. And therefore $\alpha a + (1-\alpha)c \sim b$. To justify the weaker version of the condition, we need only assume that the score is continuous on trajectories through options all of which order the worlds in the same way.

I think this justification is on the right track. But it assumes too much. We needn't assume yet that we order options by assigning a numerical score to each. Instead, we begin with $\alpha = 0$ (that is, c) and move steadily to $\alpha = 1$ (that is, a), and we ask at each point whether you prefer the mixture to b, prefer b to the mixture, are indifferent between them, or whether they are incomparable. We can, I think, assume that they are comparable. After all, a, b, and c are comparable, and it's plausible that when three options are comparable, so is any one of them with any mixture of the other two. But now suppose that there is no mixture $\alpha a + (1-\alpha)c$ for which $b \sim \alpha a + (1-\alpha)c$. Then, at some point between $\alpha = 0$ and $\alpha = 1$, my preferences just flip from preferring the mixture to preferring b; and they do this without going through the intermediate stage in which they are indifferent between the two. What could precipitate this?

One suggestion is that, at that point, the way in which the mixture orders the worlds changes. Consider, for instance, the two options $a = (1, 2)$ and $c = (2, 1)$. Then, for $\alpha < \frac{1}{2}$, the mixture $\alpha a + (1-\alpha)c$ orders the worlds strictly, with w_1 above w_2; then, at $\alpha = \frac{1}{2}$, it takes them to be equal; and, for $\alpha > \frac{1}{2}$, it orders them strictly again, with w_2 above w_1. Perhaps, then, for the risk-sensitive agent, who cares about best cases, second-best cases, ..., second-worst cases, and worst cases, we might expect a discontinuity at $\alpha = \frac{1}{2}$. We might expect preferences to flip at that point without passing through the intermediate stage. For this reason, I suspect that (A3) is a more appropriate assumption for permissivists about risk, while the stronger (A3*) is safe for others.

Note that this is the first time we appeal to our background Jamesian assumption that risk-sensitivity is permissible in the epistemic realm. If we think that is correct, we might wish to allow for the sort of discontinuity described in the previous paragraph in cases in which the two options between which we're taking mixtures order the worlds in a different way. So we opt for (A3), not (A3*).

(A4) **Weak Dominance**

(i) If $a(w) \leq a'(w)$ for all w in $\mathcal{W}$, then $a \preceq a'$.
(ii) If $a(w) < a'(w)$ for all w in $\mathcal{W}$, then $a \prec a'$.

Thus, for instance, by (ii),

$$(0, 0, \ldots, 0, 0) \prec (1, 1, \ldots, 1, 1)$$

(A4*) **Strong Dominance** If

(i) $a(w) \leq a'(w)$ for all w in $\mathcal{W}$, and

(ii) $a(w) < a'(w)$ for some w in $\mathcal{W}$,

then $a \prec a'$.

Thus, for instance,

$$(0, 1, \ldots, 1, 1) \prec (1, 1, \ldots, 1, 1)$$

The weak version of the dominance principle is uncontroversial; the strong version is slightly more controversial. After all, even standard expected decision theory sometimes allows you to violate the latter. If I assign zero probability to world w, and if there are two options, a and a', which have the same utilities everywhere except at w, and a' has greater utility than a at w, then a' weakly dominates a, but their subjective expected utility is equal—the only place they differ is at w and that receives probability weight 0.

However, even subjective Bayesians are keen to avoid this outcome in cases in which you have no evidence. They typically impose the Regularity Principle, which, in its most plausible form says that you should assign positive probability to all worlds at which your evidence is true. Your prior is your response to an absence of evidence, so your evidence at the point when you are picking priors is true at every world. And so the Regularity Principle secures a positive probability for each world, and maximising expected utility relative to such a prior will always satisfy Strong Dominance.

(A5) **Weak Linearity** If m, k are real numbers, $m > 0$, and $a \sim a'$, then $ma + k \sim ma' + k$.

(A5*) **Strong Linearity** If m, k are real numbers and $a \sim a'$, then $ma + k \sim ma' + k$.

(A6) **Weak Summation** If a, a', b, and b' order the worlds in the same way, and $a \sim a'$ and $b \sim b'$, then $a + a' \sim b + b'$.

(A6*) **Strong Summation** If $a \sim a'$ and $b \sim b'$, then $a + a' \sim b + b'$.

As with the Mixture Continuity axioms, the risk-sensitive decision-maker will want to opt for the weaker versions of these requirements. After all, they might easily be indifferent between $a = (1, 8)$ and $a' = (2, 2)$, because they are a little risk-averse and focus on the low worst case of a rather than the high best case, but then would not be indifferent between $-a = (-1, -8)$ and $-a' = (-2, -2)$, since the worst case of $-a$ is much lower than the guaranteed payout of $-a'$. So they should opt for Weak

Linearity. And we can find similar examples to motivate not going beyond Weak Summation for the risk-sensitive. For instance, here is an example that shows that the Hurwicz Criterion violates Strong Summation: Let $\lambda = \frac{1}{4}$. And let

$$a = (0, 8) \quad b = (10, 2) \quad a' = (5, -11) \quad b' = (0, -8)$$

Then

$$H^{\frac{1}{4}}(a) = \frac{1}{4}8 + \frac{3}{4}0 = 2 < 4 = \frac{1}{4}10 + \frac{3}{4}2 = H^{\frac{1}{4}}(b)$$

and

$$H^{\frac{1}{4}}(a') = \frac{1}{4}5 + \frac{3}{4}(-11) = 1 < 2 = \frac{1}{4}0 + \frac{3}{4}(-8) = H^{\frac{1}{4}}(b')$$

but

$$H^{\frac{1}{4}}(a + a') = \frac{1}{4}5 + \frac{3}{4}(-3) = \frac{1}{2} > -2 = \frac{1}{4}10 + \frac{3}{4}(-6) = H^{\frac{1}{4}}(b + b')$$

so $a \prec a'$ and $b \prec b'$, but $a + a' \succ b + b'$. We will see another case when we come to apply these principles to the choice of credences below.

(A7) **Permutation Indifference** If $\pi : W \cong W$ is a permutation of the worlds in $\mathcal{W}$ and if $a'(w) = a(\pi(w))$ for all w in $\mathcal{W}$, then $a \sim a'$.

Thus, for instance,

$$(u_1, u_2, u_3) \sim (u_1, u_3, u_2) \sim (u_2, u_1, u_3) \sim (u_2, u_3, u_1) \sim (u_3, u_1, u_2) \sim (u_3, u_2, u_1)$$

This says that it doesn't matter to which particular worlds an option assigns its utilities. What matters is which utilities it assigns. So, if you take one option and obtain another by assigning the same utilities but to different worlds, you should be indifferent between the two. For me, this is the least controversial of the axioms when our decision rule is intended to govern us in the absence of evidence and the absence of probabilities.

Of course, it is precisely such considerations that objective Bayesians often adduce to argue for the Principle of Indifference, which says that your prior credence function should be the uniform distribution—it's the only probabilistic credence function for which maximising expected utility will satisfy Permutation Indifference. And of course this gives us impermissivism about credences. But interestingly, as we will see, if we do not assume expected utility theory from the start, we will end up with a much less restrictive decision theory, and one that is equipped to deliver permissivism. This seems to me an important point: symmetry principles like Permutation Indifference have been a battleground between subjective and objective Bayesians for many years, and indeed between other sorts of permissivist and impermissivist. As we will see, the permissivist can and indeed should accept Permutation Indifference—though only for this first crucial decision where she picks her priors.

We now come to the condition that is rather different from what has gone before. Our other conditions talk only of the individual's preference ordering over the options defined on the their set $\mathcal{W}$ of possible worlds. But to state this condition, we need to talk not only of that preference ordering, but also their preference ordering over certain options defined on coarse-grainings of $\mathcal{W}$.

So, first question: what is a coarse-graining of $\mathcal{W}$? Here's an example: Suppose my set of possible worlds is $\mathcal{W} = \{w_1, w_2, w_3\}$, where Biden wins the election in w_1, Trump in w_2, and Jorgensen in w_3. Then one coarse-graining of that set would be $\mathcal{W}' = \{w'_1, w'_2\}$, where Biden wins in w'_1, and Biden loses in w'_2. We represent a coarse-graining of $\mathcal{W}$ as a set $\mathcal{W}'$ and a surjective (or onto) function $h : \mathcal{W} \to \mathcal{W}'$.[19] The idea is that each possible world w in $\mathcal{W}$ is a version of the coarser-grained possible world $h(w)$ in $\mathcal{W}'$, but specified at a finer grain of detail. Thus, in the case of the presidential candidates above, we represent this coarse-graining by the set $\mathcal{W}' = \{w'_1, w'_2\}$ and the function h, where $h(w_1) = w_1$ and $h(w_2) = h(w_3) = w'_2$. w_2 is a version of $h(w_2) = w'_2$, but specified at a finer grain of detail: w_2 specifies not only that Biden loses, but also that Trump wins.

Now, some options defined on $\mathcal{W}$ are also well defined on a coarse-graining $\mathcal{W}'$ of $\mathcal{W}$. Suppose that, for all w_1, w_2 in $\mathcal{W}$, if $h(w_1) = h(w_2)$, then $a(w_1) = a(w_2)$. Then we can define an option a' on $\mathcal{W}'$ as follows: for w in $\mathcal{W}$, $a'(h(w)) = a(w)$. We call a' a coarse-graining of a. The following condition demands that your preference ordering $\sim$ over the options defined on $\mathcal{W}$ is related to your preference ordering $\sim'$ over the coarse-grained options defined on $\mathcal{W}'$ in a particular way.

(A8) **Coarse Grain Indifference** Suppose $\mathcal{W}'$ together with $h : \mathcal{W} \to \mathcal{W}'$ is a coarse-graining of $\mathcal{W}$. Suppose option a' defined on $\mathcal{W}'$ is a coarse-graining of option a defined on $\mathcal{W}$; and suppose option b' defined on $\mathcal{W}'$ is a coarse-graining of option b defined on $\mathcal{W}$. Then $a \sim b$ iff $a' \sim' b'$.

For instance, suppose $\mathcal{W} = \{w_1, w_2, w_3\}$, $\mathcal{W}' = \{w'_1, w'_2\}$, and $h(w_1) = w'_1$ and $h(w_2) = h(w_3) = w'_2$. Then $(u, v, v) \sim (v, u, u)$ iff $(u, v) \sim' (v, u)$.

This axiom requires that you are not only indifferent to the worlds to which the utilities are assigned, but also to the grain at which the possibilities are represented in the worlds. Suppose, for instance, we consider three

[19] h is surjective (or onto) if, for all w' in $\mathcal{W}'$, there is w in $\mathcal{W}$ such that $h(w) = w'$. That is, $\mathcal{W}'$ is the range of h.

possible worlds: one in which Joe wins the election (w_1), one in which Jo does (w_2), and one in which Donald does (w_3). I consider two options: $a = (10, 2, 2)$ and $b = (2, 10, 10)$. Now, I notice that I might have faced these options even if I'd grained the possibilities more coarsely. Suppose that my set of possible worlds is not $\mathcal{W} = \{w_1, w_2, w_3\}$, but $\mathcal{W}' = \{w'_1, w'_2\}$ with $h(w_1) = w'_1$ and $h(w_2) = h(w_3) = w'_2$, so that w'_1 is the possibility in which Joe wins, and w'_2 is the possibility in which Joe loses. Then, relative to that coarse-graining, I face the choice $a' = (10, 2)$ and $b' = (2, 10)$. These are just the coarse-grained versions of a and b. So, I should be indifferent between a and b iff I am indifferent between a' and b'.

I do not assume these axiom, and indeed I think it's not a requirement of rationality. I'll explain why below.

(A9) **Convexity** If $a \sim b$, then $a, b \preceq \frac{1}{2}a + \frac{1}{2}b$

Thus, for instance, if $(u, v) \sim (u', v')$, then

$$(u, v), (u', v') \preceq \left(\frac{1}{2}(u + u'), \frac{1}{2}(v + v')\right).$$

This is very much a requirement that only a risk-averse individual would accept. Mixtures of options are more cautious than the options they mix; so many risk-inclined individuals will prefer two options between which they are indifferent to any mixture of those options, and in particular the equal mixture of them. To see this in action, consider the Hurwicz Criterion. Suppose there are just two worlds: that is, $\mathcal{W} = \{w_1, w_2\}$. And suppose the Hurwicz weight λ is greater than $\frac{1}{2}$, so that it represents a risk-inclined agent who assigns greater weight to the best case than to the worst case. Now suppose $u > v$. Then

$$H^\lambda((u, v)) = H^\lambda((v, u)) = \lambda u + (1 - \lambda)v >$$
$$\frac{1}{2}u + \frac{1}{2}v = H^\lambda\left(\left(\frac{1}{2}u + \frac{1}{2}v, \frac{1}{2}u + \frac{1}{2}v\right)\right) = H^\lambda\left(\frac{1}{2}(u, v) + \frac{1}{2}(v, u)\right)$$

Again, I do not accept this as a requirement of rationality, so I will say nothing more to motivate it.

4.3.2 The representation theorems

This completes our tour through some of the many requirements we might place on our preference ordering. We can now move to our characterization results, which borrow heavily from John Milnor's classic paper, 'Games against Nature' (Milnor, 1954).[20] We state these in the following pair of

[20]Note that these characterization results are a little different from the representation theorems we often find in decision theory. Those representation theorems seek to extract from

tables. Down the left of each, we have various decision theories: SB is subjective Bayesianism, OB is objective Bayesianism, and so on. Along the top we have various requirements. In the first table, these requirements are imposed only on the preferences $\preceq$ that order the options defined on $\mathcal{W}$. In the second, they are imposed on $\preceq$, but also on any preferences $\preceq'$ that order the coarse-grained options defined on a coarse-graining $\mathcal{W}'$ together with $h : \mathcal{W} \to \mathcal{W}'$. The core principles are: Reflexivity, Transitivity, Weak Mixture Continuity, Weak Dominance, Weak Linearity, and Weak Summation. All of our decision theories require those. We then distinguish between our various decision theories by adding further requirements: Strong Linearity, Permutation Invariance, Grain Invariance, and Convexity. Given a decision theory, we specify the axioms required to characterise it with ✓. We put ∘ if an axiom is compatible with a decision theory, though not required to characterise it, and sometimes we say what further conditions that axiom places on the decision theory. We put × if an axiom is incompatible with the decision theory. And we put ? if we just don't know!

Theorem 5

(I) *Imposing conditions on $\preceq$ only:*

		A1-6	A4*	A5*	A7	A8	A9
(i)	SB	✓	$p_i > 0$	✓	$p_i = \frac{1}{n}$	∘	∘
(ii)	OB	✓	∘	✓	✓	∘	∘
(iii)	RSB	✓	$p_i > 0$	$r(x) = x$	$p_i = \frac{1}{n}$	?	?
(iv)	ROB	✓	∘	$r(x) = x$	✓	×	?
(vii)	GHC	✓	$\lambda_i > 0$	$\lambda_i = \frac{1}{n}$	✓	×	?

(II) *Imposing conditions on $\preceq$, as well as $\preceq'$, for any coarse-graining $\mathcal{W}'$ together with $h : \mathcal{W} \to \mathcal{W}'$:*

		A1-6	A4*	A5*	A7	A8	A9
(v)	MM	✓	×	×	✓	✓	✓
(vi)	HC	✓	×	$\lambda = \frac{1}{2}$	✓	✓	∘

I hope I have gone some way towards motivating the core principles A1-6 above. And I hope I've convinced you that someone who is permissivist about attitudes to risk should be wary of Strong Summation (A5*), and the permissivist about risk who wishes to permit risk-inclined attitudes

preference orderings not only credences or risk weights, but also the utilities themselves. In our characterization, we follow Hurwicz and Milnor in assuming that the cardinal utilities are given already. This is appropriate in our case, since we are interested in the doxastic case, where the epistemic utilities are already specified by the various scoring rules.

as well as risk-averse attitudes should be wary of Convexity (A9). I wish to combine (A1-6) with Permutation Indifference (A7) to give the generalised Hurwicz Criterion. But this raises the question: why not go further and adopt Coarse Grain Indifference (A8) as well? That would give the original Hurwicz criterion. After all, both seem compelling principles for an individual at the beginning of their epistemic life, who has no probabilities to guide them. The first says that, for such an individual, it shouldn't matter which utilities are assigned to which worlds, but only which utilities are assigned and to how many worlds; the second says that, for such an individual, it should matter the level of grain at which a decision is presented.

The problem is that Coarse Grain Indifference (A8) is inconsistent with Permutation Indifference (A7) and Strong Dominance (A4*). The problem is easily seen. Consider two options (u, u, v) and (u, v, v), where $u > v$. Then, by Permutation Indifference for $\sim$, $(u, u, v) \sim (v, u, u)$. By Coarse Grain Indifference $(v, u, u) \sim (u, v, v)$ iff $(v, u) \sim' (u, v)$. By Permutation Indifference for $\sim'$, $(v, u) \sim' (u, v)$. So $(u, v, v) \sim (v, u, u) \sim (u, u, v)$. But (u, u, v) weakly dominates (u, v, v). So, by Strong Dominance, $(u, v, v) \prec (u, u, v)$. And we have a contradiction. I think this is an interesting result.[21] All three conditions—Strong Dominance, Permutation Indifference, and Coarse Grain Indifference—seem extremely plausible, and indeed fundamental, in the absence of evidence. And this is one of the times where, as I mentioned in the Introduction, we sail close to inconsistency. Here, I side with Strong Dominance and Permutation Indifference, and I reject Coarse Grain Indifference. Given the inconsistency, and the resulting forced choice between them, I think this is the right pair to back. Strong Dominance should be endorsed by anyone, at least for individuals who lack any evidence at all. And the permissivist about risk attitudes should endorse Permutation Indifference instead of Coarse Grain Indifference. A risk-sensitive agent might easily treat (u, v, v) and (u, v) differently. In the former, there is a best case, a second-best case, and a worst case; in the latter only a best and a worst. It seems reasonable that this might make a difference. On the other hand, there can be no justification for treating (u_1, u_2, u_3) differently from (u_1, u_3, u_2) in the absence of evidence.

Together with the core principles, Permutation Indifference delivers GHC. Adding Strong Dominance adds the requirement that each Hurwicz weight is positive. It is that decision rule whose consequences I will explore in the epistemic realm in the remainder of this book.

[21] After I posted a blogpost discussing it in 2020, Johan Gustafsson sent me an unpublished paper of his in which he notes the same fact (Gustafsson, ms).

4.4 Appendix: proofs

4.4.1 A useful lemma

We begin with a lemma that can be proved using only the axioms in the core, namely, Reflexivity, Transitivity, Weak Mixture Continuity, and Weak Dominance (i.e. A1-4).

Lemma 6 *If $\preceq$ satisfies (A1-4), then, for any option a, there is a real number β such that*

$$\overline{\beta} = (\beta, \dots, \beta) \sim a$$

Proof of Lemma 6. Suppose $\preceq$ satisfies Continuity and Weak Dominance. And suppose that, for all worlds w_i, $u \leq a(w_i) \leq v$. By Weak Dominance,

$$\overline{u} \preceq a \preceq \overline{v}$$

So, by Weak Mixture Continuity, there is α such that

$$\alpha\overline{u} + (1-\alpha)\overline{v} \sim a$$

But

$$\alpha\overline{u} + (1-\alpha)\overline{v} = \overline{\alpha u + (1-\alpha)v}$$

So let $\beta = \alpha u + (1-\alpha)v$. Then $\overline{\beta} \sim a$, as required. □

In the coming sections, we prove Theorem 5. We will provide our proofs specifically for the case in which there are three worlds. It is easy to generalise from these proofs, but we avoid tendrils of sub- and superscripts by doing so. Throughout, we appeal to Lemma 6 to define the following numbers:

$$\begin{array}{lcl} (1,0,0) & \sim & \overline{\beta_1} \\ (0,1,0) & \sim & \overline{\beta_2} \\ (0,0,1) & \sim & \overline{\beta_3} \end{array} \qquad \begin{array}{lcl} (1,1,0) & \sim & \overline{\beta_4} \\ (0,1,1) & \sim & \overline{\beta_5} \\ (1,0,1) & \sim & \overline{\beta_6} \end{array}$$

Note that, by Weak Dominance,

$$\begin{array}{ccccccc} 0 & \leq & \beta_1 & \leq & \beta_4, \beta_6 & \leq & 1 \\ 0 & \leq & \beta_2 & \leq & \beta_4, \beta_5 & \leq & 1 \\ 0 & \leq & \beta_3 & \leq & \beta_5, \beta_6 & \leq & 1 \end{array}$$

4.4.2 Proof of Theorem 5(i): characterizing subjective Bayesianism

Define the probability assignment P on $\mathcal{W}$ as follows:

$$\begin{array}{lcl} p_1 & = & \beta_1 \\ p_2 & = & \beta_4 - \beta_1 \\ p_3 & = & 1 - \beta_4 \end{array}$$

Then $0 \leq p_1, p_2, p_3 \leq 1$ and

$$p_1 + p_2 + p_3 = \beta_1 + (\beta_4 - \beta_1) + (1 - \beta_4) = 1$$

Then suppose $a = (u_1, u_2, u_3)$. By Strong Linearity (A5*),

$$\begin{aligned}
(\beta_1(u_1 - u_2), \beta_1(u_1 - u_2), \beta_1(u_1 - u_2)) &\sim (u_1 - u_2, 0, 0) \\
(\beta_4(u_2 - u_3), \beta_4(u_2 - u_3), \beta_4(u_2 - u_3)) &\sim (u_2 - u_3, u_2 - u_3, 0) \\
(u_3, u_3, u_3) &\sim (u_3, u_3, u_3)
\end{aligned}$$

Now,

$$\beta_1(u_1 - u_2) + \beta_4(u_2 - u_3) + u_3 = p_1 u_1 + p_2 u_2 + p_3 u_3 = \mathrm{Exp}^P(a)$$

and

$$(u_1 - u_2) + (u_2 - u_3) + u_3 = u_1$$

and

$$(u_2 - u_3) + u_3 = u_2$$

So, by Strong Summation (A6*),

$$\overline{\mathrm{Exp}^P(a)} = \overline{\sum_i p_i u_i} = \left(\sum_i p_i u_i, \sum_i p_i u_i, \sum_i p_i u_i \right) \sim (u_1, u_2, u_3)$$

So, if $\mathrm{Exp}^P(a) = \sum_i p_i u_i = \sum_i p_i v_i = \mathrm{Exp}^P(b)$, then by Transitivity (A2)

$$a = (u_1, u_2, u_3) \sim \overline{\sum_i p_i u_i} \sim \overline{\sum_i p_i v_i} \sim (v_1, v_2, v_3) = b$$

What's more, if $\mathrm{Exp}^P(a) = \sum_i p_i u_i < \sum_i p_i v_i = \mathrm{Exp}^P(b)$, then by Weak Dominance (A4) and Transitivity (A2),

$$(u_1, u_2, u_3) \sim \overline{\sum_i p_i u_i} \prec \overline{\sum_i p_i v_i} \sim (v_1, v_2, v_3)$$

Thus, $a \preceq b$ iff $\mathrm{Exp}^P(a) \preceq \mathrm{Exp}^P(b)$, as required. □

4.4.3 Proof of Theorem 5(ii): characterizing objective Bayesianism

By Theorem 5(i), there is a probability assignment P such that $a \preceq b$ iff $\sum_i p_i a(w_i) \preceq \sum_i p_i b(w_i)$. By Permutation Indifference (A7), $p_1 = p_2 = p_3$, as required. □

4.4.4 Proof of Theorem 5(iii): characterizing subjective risk-weighted Bayesianism

Let

$$\begin{array}{rclcrcl} r(p_1) & = & \beta_1 & \quad & r(p_1+p_2) & = & \beta_4 \\ r(p_2) & = & \beta_2 & & r(p_2+p_3) & = & \beta_5 \\ r(p_3) & = & \beta_3 & & r(p_1+p_3) & = & \beta_6 \end{array}$$

Then suppose $a = (u_1, u_2, u_3)$. And suppose that $u_1 \geq u_2 \geq u_3$. Then, by Weak Linearity (A5),

$$\begin{array}{rcl} (\beta_1(u_1-u_2), \beta_1(u_1-u_2), \beta_1(u_1-u_2)) & \sim & (u_1-u_2, 0, 0) \\ (\beta_4(u_2-u_3), \beta_4(u_2-u_3), \beta_4(u_2-u_3)) & \sim & (u_2-u_3, u_2-u_3, 0) \\ (u_3, u_3, u_3) & \sim & (u_3, u_3, u_3) \end{array}$$

Now,

$$\beta_1(u_1-u_2) + \beta_4(u_2-u_3) + u_3 = \\ r(p_1)(u_1-u_2) + r(p_1+p_2)(u_2-u_3) + u_3 = \mathrm{RExp}^{P,r}(a)$$

and

$$(u_1-u_2) + (u_2-u_3) + u_3 = u_1$$

and

$$(u_2-u_3) + u_3 = u_2$$

So, by Weak Summation (A6),

$$\overline{\mathrm{RExp}^{P,r}(a)} \sim (u_1, u_2, u_3)$$

So, if $\mathrm{RExp}^{P,r}(a) = \mathrm{RExp}^{P,r}(b)$, then, by Transitivity (A2), $a \sim b$. What's more, if $\mathrm{RExp}^{P,r}(a) < \mathrm{RExp}^{P,r}(b)$, then by Weak Dominance (A4) and Transitivity (A2), $a \prec b$. Thus, $a \preceq b$ iff $\mathrm{RExp}^{P,r}(a) \leq \mathrm{RExp}^{P,r}(b)$, as required. □

4.4.5 Proof of Theorem 5(iv): characterizing objective risk-weighted Bayesianism

By Theorem 5(iii), there is a probability assignment P and a risk function r such that $a \preceq b$ iff $\mathrm{RExp}^{P,r}(a) \leq \mathrm{RExp}^{P,r}(b)$. By Permutation Indifference (A7), $p_1 = p_2 = p_3$, as required. □

4.4.6 Proof of Theorem 5(vi): characterizing the Hurwicz Criterion

Let $\lambda = \beta_1$. So

$$(\lambda, \lambda, \lambda) \sim (1, 0, 0)$$

Now, for any $u < v$, by Weak Linearity (A5), with $m = v - u > 0$ and $k = u$,

$$(u + \lambda(v - u), u + \lambda(v - u), u + \lambda(v - u)) \sim ((v - u) + u, u, u)$$

So

$$(\lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u) \sim (v, u, u)$$

By Permutation Indifference (A7) and Coarse Grain Indifference (A8),

$$(v, u, u) \sim (u, v, v) \sim (v, v, u)$$

Thus, by Transitivity (A2),

$$(\lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u) \sim (u, v, v)$$

Now, suppose $v = u_1 \geq u_2 \geq u_3 = u$. Then, by Weak Dominance (A4),

$$(u_1, u_2, u_3) \preceq (v, v, u) \sim (\lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u)$$

And, again by Weak Dominance (A4),

$$(\lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u) \sim (u, u, v) \preceq (u_1, u_2, u_3)$$

So, by Transitivity (A2),

$$(u_1, u_2, u_3) \sim (\lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u, \lambda v + (1 - \lambda)u) = \overline{H^\lambda(a)}$$

Thus, $a \preceq b$ iff $H^\lambda(a) \leq H^\lambda(b)$. □

4.4.7 Proof of Theorem 5(v): characterizing Minimax

By Theorem 5(vi), there is $0 \leq \lambda \leq 1$ such that $a \preceq b$ iff $H^\lambda(a) \leq H^\lambda(b)$. Then, note that, by Permutation Indifference,

$$(0, 0, 1) \sim (0, 1, 0)$$

So, by Convexity (A8),

$$(0, 1, 0), (0, 0, 1) \preceq (0, \frac{1}{2}, \frac{1}{2})$$

But

- $H^\lambda((0, 0, 1)) = \lambda$
- $H^\lambda((0, \frac{1}{2}, \frac{1}{2})) = \frac{1}{2}\lambda$

So $H^\lambda((0, 0, 1)) \leq H^\lambda((0, \frac{1}{2}, \frac{1}{2}))$ only if $\lambda = 0$. But $H^0(a) = W(a)$, so $a \preceq b$ iff $\mathcal{W}(a) \leq W(b)$, as required. □

4.4.8 Proofs of Theorems 5(vii): characterizing the generalised Hurwicz Criterion

Let

$$\begin{aligned}\lambda_1 &= \beta_1\\ \lambda_2 &= \beta_4 - \beta_1\\ \lambda_3 &= 1-(\beta_4-\beta_1)-\beta_1 = 1-\beta_4\end{aligned}$$

Then $0 \leq \lambda_1, \lambda_2, \lambda_3 \leq 1$ and $\lambda_1+\lambda_2+\lambda_3 = 1$. Also,

$$\begin{aligned}(\lambda_1,\lambda_1,\lambda_1) &\sim (1,0,0)\\ (\lambda_1+\lambda_2,\lambda_1+\lambda_2,\lambda_1+\lambda_2) &\sim (1,1,0)\\ (\lambda_1+\lambda_2+\lambda_3,\lambda_1+\lambda_2+\lambda_3,\lambda_1+\lambda_2+\lambda_3) &\sim (1,1,1)\end{aligned}$$

Now, suppose $u_1 \geq u_2 \geq u_3$, and consider the option $a = (u_1, u_2, u_3)$. Then, by Weak Linearity (A5):

- $(\lambda_1(u_1-u_2), \lambda_1(u_1-u_2), \lambda_1(u_1-u_2)) \sim (u_1-u_2, 0, 0)$
- $((\lambda_1+\lambda_2)(u_2-u_3), (\lambda_1+\lambda_2)(u_2-u_3), (\lambda_1+\lambda_2)(u_2-u_3)) \sim (u_2-u_3, u_2-u_3, 0)$
- $((\lambda_1+\lambda_2+\lambda_3)u_3, (\lambda_1+\lambda_2+\lambda_3)u_3, (\lambda_1+\lambda_2+\lambda_3)u_3) \sim (u_3, u_3, u_3)$

And so, by repeated application of Weak Summation (A6) and Transitivity (A2):

$$\overline{H^{\Lambda}(a)} = \left(\sum_i \lambda_i u_i, \sum_i \lambda_i u_i, \sum_i \lambda_i u_i\right) \sim (u_1, u_2, u_3)$$

So, $a \preceq b$ iff $H^{\Lambda}(a) \leq H^{\Lambda}(b)$, as required.

Now, suppose Strong Dominance (A4*). And suppose $\lambda_k = 0$. Then define options a and b as follows:

	w_1	$\dots$	w_{k-1}	w_k	w_{k+1}	$\dots$	w_n
a	1	$\dots$	1	0	0	$\dots$	0
b	1	$\dots$	1	1	0	$\dots$	0

Then, by Strong Dominance, $a \prec b$. But

$$H^{\Lambda}(a) = \lambda_1 1 + \ldots + \lambda_{k-1}1 + \lambda_k 0 + \lambda_{k+1}0 + \ldots + \lambda_n 0 = \lambda_1 + \ldots + \lambda_{k-1}$$

and

$$H^{\Lambda}(b) = \lambda_1 1 + \ldots + \lambda_{k-1}1 + \lambda_k 1 + \lambda_{k+1}0 + \ldots + \lambda_n 0 = \lambda_1 + \ldots + \lambda_{k-1} + \lambda_k = \lambda_1 + \ldots + \lambda_{k-1}$$

So $H^{\Lambda}(a) = H^{\Lambda}(b)$. And thus $a \sim b$. This gives a contradiction. So, if we assume Strong Dominance, we must have $0 < \lambda_1, \ldots, \lambda_n < 1$. □

In brief...

In this chapter, we argued in favour of the family of decision rules that we will apply to the choice of prior credences in what follows. The family is called *the generalised Hurwicz Criterion* (*GHC*). Each decision rule within the family is specified by a set of *generalised Hurwicz weights*. Each weight is a number. The first weight is the one we will apply to the best-case utility of an option to determine its contribution to the overall value of that option; the second weight is the one we will apply to the second-best case utility to determine its contribution; and so on down to the second-last weight, which is the one we will apply to the second-worst case utility, and the last weight, which we apply to the worst-case utility. We measure the value of each option by its *generalised Hurwicz score*, which we calculate as follows: we line up the possible utilities that the option might obtain for us, from best to worst. And then we apply the appropriate weights to each of these utilities—our first generalised Hurwicz weight to the best utility, second weight to the second-best utility, and so on. And then we add up the weighted utilities. We then prefer one option to another just in case the generalised Hurwicz score of the first exceeds the generalised Hurwicz score of the second. And we say that an option is irrational if there is some alternative that we strictly prefer. The result is that the rationally permissible options relative to this decision rule are precisely those that maximise this generalised Hurwicz score.

We might understand GHC by contrasting it with a more familiar family of decision rules, namely, the family of standard expected utility rules. According to the rules in both families, the utility of an option at a world receives a particular weighting, and you score an option by applying the appropriate weightings and adding up the weighted utilities. But they differ in the way in which the weights are assigned. In expected utility, each world receives a weight and that weight is applied to the utility of the option at that world. In GHC, in contrast, a weight does not attach to a specific world, but to a position in the ranking of worlds from best to worst. That weight is then applied to the utility of the option at the world that occupies that place in the ranking of worlds by the utilities that the option obtains at those worlds.

These generalised Hurwicz weights are best understood as our attitudes to risk. The more risk-inclined you are, the more weight you will concentrate in the earlier generalised Hurwicz weights, since these weight the better cases; the more risk-averse, the more weight you will concentrate in the latter generalised Hurwicz weights, since these weight the worse cases. And if you're risk-neutral, you'll assign equal generalised Hurwicz weights to each position in the ranking of worlds.

5 Epistemic risk and picking priors II: the consequences of the rule

The upshot of Section 3.2 was that we cannot argue for permissivism about credences in exactly the same way that Kelly argues for permissivism about full beliefs. That is, while there are many permissible measures of epistemic value for credences, just as there are for full beliefs, the credal epistemic utility functions all agree on which credence maximises expected epistemic value from the point of view of a specific evidential probability—it is that very evidential probability itself. So, if we are to establish permissivism by appealing to the notion of epistemic risk, we can't do as Kelly does and encode attitudes to epistemic risk in our measures of epistemic value. We must instead encode them in our decision rules. That is, we must use one of the decision rules discussed in Section 4.2. For the reasons given there, when it is our prior credences that we are picking, we'll use the generalised Hurwicz Criterion.

So, when we pick our priors, we encode the way we value credences in our epistemic value measure $\mathfrak{EU}$; and we encode our attitudes to risk in our generalised Hurwicz weights, $0 \leq \lambda_1, \ldots, \lambda_n \leq 1$; and then we pick a prior that maximises the generalised Hurwicz score with these weights. That is, the norm that governs our priors is this:

> **Rational Priors** If $\mathfrak{EU}$ is your epistemic utility function for credences functions, then rationality requires you to pick your generalised Hurwicz weights, $0 \leq \lambda_1, \ldots, \lambda_n \leq 1$, which we'll take to encode your attitudes to epistemic risk, and then to pick a prior C with maximal generalised Hurwicz score relative to those weights and that epistemic utility function.[22]
>
> That is, you are rationally permitted to pick C if, and only if, for any credence function C',
>
> $$H^{\Lambda}(\mathfrak{EU}(C)) \geq H^{\Lambda}(\mathfrak{EU}(C'))$$
>
> where $\mathfrak{EU}(C)$ is the option whose utility at world w is $\mathfrak{EU}(C, w)$, and similarly for $\mathfrak{EU}(C')$.
>
> That is, you are rationally permitted to pick C if, and only if, for

[22] I should perhaps emphasise that this description of how you should go about setting your prior is not intended to be taken too literally. GHC doesn't demand that you go through the process described here. It simply says that your preference ordering over credence functions should be representable by a generalised Hurwicz score. That is, there should be generalised Hurwicz weights such that you prefer one credence function to another just in case the generalised Hurwicz score of the first relative to those weights and relative to your epistemic utility function is greater than the generalised Hurwicz score of the second.

any credence function C', if

$$\mathfrak{EU}(C, w_{i_1}) \geq \ldots \geq \mathfrak{EU}(C, w_{i_n})$$

and

$$\mathfrak{EU}(C', w_{j_1}) \geq \ldots \geq \mathfrak{EU}(C', w_{j_n})$$

Then

$$H^{\Lambda}(\mathfrak{EU}(C)) = \lambda_1 \mathfrak{EU}(C, w_{i_1}) + \ldots + \lambda_n \mathfrak{EU}(C, w_{i_n}) \geq \lambda_1 \mathfrak{EU}(C', w_{j_1}) + \ldots + \lambda_n \mathfrak{EU}(C', w_{j_n}) = H^{\Lambda}(\mathfrak{EU}(C'))$$

5.1 Decomposing options into their component parts

In the next section, we'll explore the consequences of this norm. But first, we should note something about our decision to apply GHC to our choice of prior credence function, rather than to our choice of individual prior credences. As we're about to see, this is a substantial choice.

Consider the simplest of individuals. They have credences in just two possible worlds, w_1 and w_2. And they assign credences to w_1 or w_2 or both. And suppose that their generalised Hurwicz weights are $\lambda_1 = \frac{3}{4}$ and $\lambda_2 = \frac{1}{4}$.[23]

Now, first of all, apply GHC with these weights to the choice of credence function. As always, we measure the epistemic value of a whole credence function using a strictly proper measure, $\mathfrak{EU}$, generated by a strictly proper scoring rule, $\mathfrak{s}$. Then it turns out that exactly two credence functions maximise $H^{\left(\frac{3}{4}, \frac{1}{4}\right)}(\mathfrak{EU}(-))$. They are:

	w_1	w_2
C_1	3/4	1/4
C_2	1/4	3/4

Next, apply GHC with these weights to the choice of individual credences when we measure the epistemic value of an individual credence with the strictly proper scoring rule, $\mathfrak{s}$. There are two credences you might assign to w_1 that maximise $H^{\left(\frac{3}{4}, \frac{1}{4}\right)}(\mathfrak{s}(-))$. They are: $\frac{1}{4}$ and $\frac{3}{4}$. And there are two credences you might assign to w_1 that maximise $H^{\left(\frac{3}{4}, \frac{1}{4}\right)}(\mathfrak{s}(-))$. Again, they are: $\frac{1}{4}$ and $\frac{3}{4}$. So, if we apply GHC when we pick our prior credence in w_1, it's rationally permissible to assign it $\frac{1}{4}$; and if we apply it when we pick our prior in w_2, it is again rationally permissible to assign $\frac{1}{4}$ to that as well.

[23]Note that, since there are only two worlds in play, applying GHC with $\lambda_1 = \frac{3}{4}$ and $\lambda_2 = \frac{1}{4}$ is equivalent to applying Hurwicz's original criterion with $\lambda = \frac{3}{4}$.

But of course this results in a credence function C', which is both different from C_1 and C_2 and also non-probabilistic:

	w_1	w_2
C'	$1/4$	$1/4$

This is a particular instance of the phenomenon we met above: GHC does not obey the Strong Summation (A6*) axiom from above. Suppose I measure epistemic utility using what we sometimes call a $0/1$ *symmetric scoring rule*: that is, $\mathfrak{s}(1, p) = \mathfrak{s}(0, 1-p)$, for all $0 \leq p \leq 1$—the quadratic score $\mathfrak{q}$ is an example of such a scoring rule. Then, if I assign $\frac{1}{4}$ to w_1 (and no credence to w_2), that receives exactly the same generalised Hurwicz score as if I assign $\frac{3}{4}$ to w_1 (and no credence to w_2). And, obviously, if I assign $\frac{1}{4}$ to w_2 (and no credence to w_1), that receives the same generalised Hurwicz score as if I assign $\frac{1}{4}$ to w_2 (and no credence to w_1). Yet, if I assign $\frac{1}{4}$ to w_1 *and* $\frac{1}{4}$ to w_2, I do not receive the same generalised Hurwicz score as if I assign $\frac{3}{4}$ to w_1 *and* $\frac{1}{4}$ to w_2—indeed, I receive a lower score.[24]

The upshot is that GHC is sensitive to whether an option is presented to you all at once or in parts spread across different decision problems. If there is a way to break options down into component parts, you might hope that your decision rule does not make different demands depending on whether you are faced with a choice between two whole options and a choice between each of their component parts separately. It is well known that risk-sensitive decision rules typically violate this requirement (Buchak, 2013, Chapters 6-7). It is also well known that this often leads them to be exploitable in various ways. That is, they will permit you to choose a dominated option when it and the alternatives are presented piecemeal.[25] And

[24]The epistemic utility of assigning $\frac{1}{4}$ to w_1 *and* $\frac{1}{4}$ to w_2 is the same at the two worlds: it is $\mathfrak{s}(0, \frac{1}{4}) + \mathfrak{s}(1, \frac{1}{4}) = \mathfrak{s}(1, \frac{3}{4}) + \mathfrak{s}(0, \frac{3}{4})$. So that is its generalised Hurwicz score.

The epistemic utility of assigning $\frac{3}{4}$ to w_1 *and* $\frac{1}{4}$ to w_2 is $\mathfrak{s}(1, \frac{3}{4}) + \mathfrak{s}(0, \frac{1}{4})$ at w_1 and $\mathfrak{s}(0, \frac{3}{4}) + \mathfrak{s}(1, \frac{1}{4})$ at w_2. Since, $\mathfrak{s}(1, \frac{3}{4}) + \mathfrak{s}(0, \frac{1}{4}) > \mathfrak{s}(0, \frac{3}{4}) + \mathfrak{s}(1, \frac{1}{4})$, its generalised Hurwicz score is:

$$\frac{3}{4}\left(\mathfrak{s}(1, 3/4) + \mathfrak{s}(0, 1/4)\right) + \frac{1}{4}\left(\mathfrak{s}(0, 3/4) + \mathfrak{s}(1, 1/4)\right)$$

Since $\mathfrak{s}$ is strictly proper,

$$\begin{aligned}
&\frac{3}{4}\left(\mathfrak{s}(1, 3/4) + \mathfrak{s}(0, 1/4)\right) + \frac{1}{4}\left(\mathfrak{s}(0, 3/4) + \mathfrak{s}(1, 1/4)\right) = \\
&\qquad \left(\frac{3}{4}\mathfrak{s}(1, 3/4) + \frac{1}{4}\mathfrak{s}(0, 3/4)\right) + \left(\frac{1}{4}\mathfrak{s}(1, 1/4) + \frac{3}{4}\mathfrak{s}(0, 1/4)\right) > \\
&\qquad \left(\frac{3}{4}\mathfrak{s}(1, 1/4) + \frac{1}{4}\mathfrak{s}(0, 1/4)\right) + \left(\frac{1}{4}\mathfrak{s}(1, 1/4) + \frac{3}{4}\mathfrak{s}(0, 1/4)\right) = \\
&\qquad\qquad \mathfrak{s}(0, 1/4) + \mathfrak{s}(1, 1/4) = \mathfrak{s}(1, 3/4) + \mathfrak{s}(0, 3/4)
\end{aligned}$$

[25]Indeed, unless your generalised Hurwicz weights are all the same—in which case GHC

indeed that's exactly what happens in the case I have just presented. Assigning credence $\frac{1}{4}$ to w_1 and credence $\frac{1}{4}$ to world w_2 is permissible if your choice of credence for each world is presented separately. But, since this option is non-probabilistic, we know that it is dominated. And indeed, for Maximin, the Hurwicz Criterion, and the generalised Hurwicz Criterion, there are dominated options that are not only permissible when presented individually, but mandated. For instance, consider the following three options:

	w_1	w_2
a	0	8
b	8	0
c	3	3

And suppose you're first presented with the choice between a and c, and then with the choice between b and c. Maximin will lead you to choose c both times. But $c + c$ is dominated by $a + b$. What's more, the Hurwicz criterion with $\lambda < \frac{3}{8}$ will lead to the same choices; and similarly for the generalised Hurwicz Criterion with $\lambda_1 < \frac{3}{8}$.

One way to avoid this issue is to strengthen our axioms above by imposing Strong Summation (axiom A6*) instead of merely Weak Summation (axiom A6). Along with the other axioms for which we argued in that section, this would give us objective Bayesianism and, with it, impermissivism. But to do so would be to say that different attitudes to risk are not permissible. After all, objective Bayesianism does not permit them. Perhaps that's the

---

coincides with objective Bayesianism—there will always be a sequence of decision problems such that applying GHC with your weights when faced with each requires you to choose a sequence of options that is dominated by an alternative sequence. After all, suppose $\lambda_1, \ldots, \lambda_n$ is a sequence of Hurwicz weights, with $\sum_{i=1}^{n} \lambda_i = 1$. And suppose that, for some $1 \leq i \leq n$, $\lambda_i \neq \frac{1}{n}$. Then *either*

(a) there is a sequence of utilities $u_1 \geq u_2 \geq \ldots \geq u_n$ and a utility v such that

  (i) $\lambda_1 u_1 + \ldots + \lambda_n u_n > v$ and

  (ii) $\frac{u_1 + \ldots + u_n}{n} < v$

or

(b) there is a sequence of utilities $u_1 \geq u_2 \geq \ldots \geq u_n$ and a utility v such that

  (iii) $\lambda_1 u_1 + \ldots + \lambda_n u_n < v$ and

  (iv) $\frac{u_1 + \ldots + u_n}{n} > v$

Suppose (a). Then, for any permutation π of the worlds $w_1, \ldots, w_n$, consider the following options: $u_\pi = (u_{\pi(1)}, \ldots, u_{\pi(n)})$ and $v = (v, \ldots, v)$. By (i), GHC requires you to choose u_π over v. Now suppose that you face $n!$ decision problems, one for each permutation π. In the decision problem corresponding to permutation π, the options are u_π and v. In each, GHC requires you to choose u_π. Then the total utility of your sequence of choices at each world is $(n-1)!(u_1 + \ldots + u_n)$. But the total utility of instead choosing v in each case is $n!v$ at each world. But, by (ii), $n!v > (n-1)!(u_1 + \ldots + u_n)$. Thus, the sequence of choices that GHC requires you to make is dominated. And similarly, mutatis mutandis, for (b).

correct conclusion to draw. After all, being permissive about attitudes to risk has led us to a decision theory that permits sequences of choices that, taken together, are dominated by alternative sequences of choices.

There is, I think, an alternative way out. The problem that we have identified arises if you use the GHC for more than one decision. Since GHC itself satisfies Weak Dominance—and also satisfies Strong Dominance, if the weights are all positive—then if you only ever apply the rule to a single decision, it will never permit you to choose a dominated option. And this, I suggest, is exactly what we should do. We use GHC to pick our whole credence function at the beginning of our epistemic life. That's it. That's the single decision we make using that rule. Thereafter, we are armed with probabilities, namely, our credences, and so another rule is appropriate. And, for that initial decision, we choose our whole credence function at once, for only by doing that do we guarantee that we won't pick a dominated option.

I think it's helpful to see the situation in the same way we saw the conflict between Permutation Indifference, Coarse Grain Indifference, and Strong Dominance above. Each of those three principles is compelling and desirable. But they are incompatible; they form an inconsistent triad. And so we opted for Permutation Indifference and Strong Dominance, since they are most central to the permissivism about attitudes to risk that is most central here. Similarly, we'd like our decision theories not to permit us to choose a dominated sequence of options in response to a sequence of decision problems; and we'd like to use our decision theory whenever we face a decision problem. But, we discovered here, these two properties are not compatible with being permissive about our attitudes to risk: there are sequences of decision problems such that, if we were to approach each with GHC, we would be permitted to choose a dominated sequence of options. Again, we face an inconsistent triad—and again we see what I anticipated in the Introduction, namely, that it is difficult to formulate a fully satisfactory theory that is permissive about epistemic risk in the way that James would like; inconsistency or impermissivism always threatens. Again, we pick the two that allow us to retain permissivism about attitudes to risk. That is, we retain GHC and the principle that we should not be permitted to choose a sequence of dominated options. This forces us to specify just one decision to which we apply GHC. It is our choice of prior.

5.2 The credal consequences of the rules

We are now finally at the point where we can say which priors rationality requires you to pick. We have spelled out the norm that governs this choice: it is Rational Priors. All that remains is to investigate its consequences. I'd like to begin by looking at the case in which our priors will be defined on just three possible worlds w_1, w_2, w_3. Things get complicated pretty fast as we increase the number of worlds, and there will be plenty of interest in this simple case. What's more, this is the easiest case to visualise. But we will return to the general case at the end of the next section.

We assume that our measure $\mathfrak{EU}$ of epistemic value is generated by a scoring rule in the way outlined above. That is, there is $\mathfrak{s}$ such that

$$\mathfrak{EU}(C, w) = \sum_{X \in \mathcal{F}} \mathfrak{s}(w(X), C(X))$$

And we assume that $\mathfrak{EU}$ and $\mathfrak{s}$ are strictly proper. Given a probabilistic credence function C defined on w_1, w_2, w_3, we will write c_1 for $C(w_1)$, c_2 for $C(w_2)$, and c_3 for $C(w_3)$, and we will sometimes denote C by the triple (c_1, c_2, c_3).

5.2.1 The credal consequences of the other rules

Before we turn to GHC and Rational Priors, let's see first what subjective and objective Bayesianism, Maximin, and the original Hurwicz criterion demand.

For subjective Bayesianism, it is straightforward. It permits you to pick any probability function, and then maximise expected epistemic value from that point of view. But, since $\mathfrak{s}$ and $\mathfrak{EU}$ are strictly proper, if you pick P, then P will maximise your expected epistemic utility. So every probabilistic prior is permissible from an interpersonal point of view. From an intrapersonal point of view, only the one that you pick to do the picking is permissible.

Objective Bayesian seems at first like an impermissive rule. It tells you to maximise expected epistemic value from the point of view of the uniform distribution $P^\dagger$; and since $\mathfrak{EU}$ is strictly proper scoring rule, it demands that you have the uniform distribution as your prior. In fact, as many have pointed out, this is more permissive than it seems at first, because the uniform distribution is defined relative to a set of possible worlds, and the relevant set of possible worlds is determined by the propositions about which you have opinions. And that, we might assume, can vary from rational person to rational person.

Maximin gives the same result as objective Bayesianism, though for different reasons (Pettigrew, 2016b). There is just one credence function whose worst-case is best, and it is the uniform distribution. To see this, note first that it has the same epistemic value at every world. That epistemic value is therefore both its worst and its best case. Now, consider another credence function. Since the measure is strictly proper, there can be no credence function that is at least as good as the uniform distribution at all worlds. Therefore, there is some world at which this other credence function is worse than the uniform one. But then of course its worst case is not as good as the worst case of the uniform distribution.

The original Hurwicz criterion demands that we pick a Hurwicz weight $0 \leq \lambda \leq 1$, weight the best case by λ and the worst by $1 - \lambda$, and maximise the result. In earlier work, I investigated what this decision rule requires in the credal case (Pettigrew, 2016c). There are two cases: if n is the number of possible worlds, then

(a) if $\lambda \leq \frac{1}{n}$, then the uniform distribution maximises the Hurwicz score $H^\lambda(\mathfrak{EU}(-))$.

(b) if $\lambda > \frac{1}{n}$, then there are n different credence functions that maximise the Hurwicz score $H^\lambda(\mathfrak{EU}(-))$: each assigns λ to some world, and then distributes the remaining credence $1 - \lambda$ equally over the remaining worlds, giving $\frac{1-\lambda}{n-1}$ to each.

So suppose, for instance, that there are just three possible worlds w_1, w_2, w_3. Then:

(a) if $\lambda = \frac{1}{4}$, then the uniform distribution $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ maximises the Hurwicz score;

(b) if $\lambda = \frac{3}{4}$, then the following credence functions all maximise that score:

$$\left(\frac{3}{4}, \frac{1}{8}, \frac{1}{8}\right), \quad \left(\frac{1}{8}, \frac{3}{4}, \frac{1}{8}\right), \quad \left(\frac{1}{8}, \frac{1}{8}, \frac{3}{4}\right)$$

Maximin is the most extremely risk-averse decision rule. And it demands the uniform distribution. Our characterization of the credence functions that maximise the original Hurwicz score show that you don't have to be so maximally risk-averse in order to end up with the uniform distribution. You just have to give low enough weight to the best case ($\lambda \leq \frac{1}{n}$) and high enough weight to the worst case ($1 - \lambda \geq \frac{n-1}{n}$). But if you are less risk-averse than this, you will move beyond the uniform distribution. And when you do, you will plump for a particular world and assign it greater credence than the others, and you will treat each of the others equally. As we'll see, the space of attitudes to risk that demand the uniform distribution gets even more varied when we look to the generalised Hurwicz Criterion. And, equally, the sorts of credence function that go beyond the uniform distribution that your risk attitudes might justify also become more varied.

5.2.2 The credal consequences of GHC

Having seen each of these, let's turn now to the generalised Hurwicz Criterion. Since there are three worlds, we need three generalised Hurwicz weights, $0 \leq \lambda_1, \lambda_2, \lambda_3 \leq 1$ such that $\lambda_1 + \lambda_2 + \lambda_3 = 1$. The first thing to note is that, if (c_1, c_2, c_3) maximises $H^\Lambda(\mathfrak{EU}(-))$, then so does any permutation of it—that is, the six credence functions

$$(c_1, c_2, c_3),\ (c_1, c_3, c_2),\ (c_2, c_1, c_3),\ (c_2, c_3, c_1),\ (c_3, c_1, c_2),\ (c_3, c_2, c_1)$$

all maximise $H^\Lambda(\mathfrak{EU}(-))$ if any of one of them does. The reason is that the generalised Hurwicz score for the three-world case depends on the best, middle, and worst epistemic utilities that a credence function obtains, and those are exactly the same for those six credence functions, even though they occur at different worlds for each. In other words, GHC satisfies the Permutation Indifference principle that we laid out above. This means that, in order to find the maximisers, we only need to identify the maximiser for which $c_1 \geq c_2 \geq c_3$. All others will be permutations of those. Let $\mathfrak{X} = \{(c_1, c_2, c_3) \mid c_1 \geq c_2 \geq c_3\}$. Since the measure $\mathfrak{EU}$ of epistemic value is strictly proper and generated by a scoring rule, for each c in $\mathfrak{X}$,[26]

$$\mathfrak{EU}(c, w_1) \geq \mathfrak{EU}(c, w_2) \geq \mathfrak{EU}(c, w_3)$$

And so

$$H^\Lambda(\mathfrak{EU}(c)) = \lambda_1 \mathfrak{EU}(c, w_1) + \lambda_2 \mathfrak{EU}(c, w_2) + \lambda_3 \mathfrak{EU}(c, w_3)$$

That means that $H^\Lambda(\mathfrak{EU}(c))$ is the expected inaccuracy of c by the lights of the credence function $(\lambda_1, \lambda_2, \lambda_3)$ that corresponds to the Hurwicz weights. This allows us to calculate each case. As Catrin Campbell-Moore helped me to see, it turns out that the maximiser does not depend on which strictly proper scoring rule you use to generate your measures of epistemic value—each gives the same.

In the second column of the table below, I list the different possible orderings of the three Hurwicz weights. In two cases, specifying that order is not sufficient to determine the maximiser. To do that, you also need to know the absolute values of some of the weights. Where necessary, I include those in the third column. In the fourth column, I specify the member of $\mathfrak{X}$ that maximises $H^\Lambda(\mathfrak{EU}(-))$ relative to those weights. As we noted above, any permutation of the credence function specified in the third column will also maximise $H^\Lambda(\mathfrak{EU}(-))$. In the first column, I give the region of the simplex depicted in Figure 6 that corresponds to that row of the ta-

[26]This follows from the fact that strictly proper scoring rules are truth-directed. That is, if $\mathfrak{s}$ is strictly proper, then for any $0 \leq p < q \leq 1$, $\mathfrak{s}(0, p) \geq \mathfrak{s}(0, q)$ and $\mathfrak{s}(1, q) \geq \mathfrak{s}(1, p)$. I learned this first from Catrin Campbell-Moore and Ben Levinstein. To prove it, you can use Savage's original representation theorem for strictly proper scoring rules. This says that, if $\mathfrak{s}$ is a continuous strictly proper scoring rule, then there is a strictly convex function φ such that $\mathfrak{s}(i, x) = \varphi(i) - \varphi(x) - \varphi'(x)(i - x)$.

ble.[27]

Region of simplex	Ordering of the weights	Further properties of the weights	c_1	c_2	c_3
(1)	$\lambda_1 \leq \lambda_2 \leq \lambda_3$	–	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$(2a)$	$\lambda_1 \leq \lambda_3 \leq \lambda_2$	$\lambda_3 \leq \frac{1}{3}$	$\frac{\lambda_1+\lambda_2}{2}$	$\frac{\lambda_1+\lambda_2}{2}$	λ_3
$(2b)$	$\lambda_1 \leq \lambda_3 \leq \lambda_2$	$\lambda_3 \geq \frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$(3a)$	$\lambda_2 \leq \lambda_1 \leq \lambda_3$	$\lambda_1 \leq \frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$(3b)$	$\lambda_2 \leq \lambda_1 \leq \lambda_3$	$\lambda_1 \geq \frac{1}{3}$	λ_1	$\frac{\lambda_2+\lambda_3}{2}$	$\frac{\lambda_2+\lambda_3}{2}$
(4)	$\lambda_2 \leq \lambda_3 \leq \lambda_1$	–	λ_1	$\frac{\lambda_2+\lambda_3}{2}$	$\frac{\lambda_2+\lambda_3}{2}$
(5)	$\lambda_3 \leq \lambda_1 \leq \lambda_2$	–	$\frac{\lambda_1+\lambda_2}{2}$	$\frac{\lambda_1+\lambda_2}{2}$	λ_3
(6)	$\lambda_3 \leq \lambda_2 \leq \lambda_1$	–	λ_1	λ_2	λ_3

In Figure 6, we plot the different possible generalised Hurwicz weights in a barycentric plot, so that the bottom left corner of the triangle is $(1, 0, 0)$, the bottom-right is $(0, 1, 0)$ and the top is $(0, 0, 1)$. We then divide this into four regions. If your weights $\Lambda = (\lambda_1, \lambda_2, \lambda_3)$ lie in a given region, then the triple I've placed in that region gives the credence function that maximises the generalised Hurwicz score $H^\Lambda(\mathfrak{EU}(-))$ for those weights. Note, the bottom left triangle is $\mathfrak{X}$. Essentially, to find which member of $\mathfrak{X}$ a given weighting demands, you plot that weighting in this diagram and then find the closest member of $\mathfrak{X}$, where the measure of distance is Euclidean. As mentioned above, it turns out that this will work for any measure $\mathfrak{EU}$ of epistemic value that is generated by a strictly proper scoring rule.

Let's work through the segments, from (1) through to (6). Let's start with (1). Perhaps unsurprisingly, if you assign greatest weight to the worst-case, next greatest to the middle, and least weight to the best case, GHC demands that you pick the uniform distribution, just as Maximin also demands. But, those aren't the only weights that demand the uniform distribution, as witnessed by (2b) and (3a). Notice that in all three—that is, (1), (2b), and (3a)—the weight given to the best case is at most one-third—that is, $\lambda_1 \leq \frac{1}{3}$; this mirrors the situation for the original Hurwicz criterion. However, in the original version, assigning at most than one-third to the

[27]In an n-dimensional space, the standard simplex is the set of points $P = (p_1, \ldots, p_n)$ such that $0 \leq p_1, \ldots, p_n \leq 1$ and $\sum_{i=1}^{n} p_i = 1$.

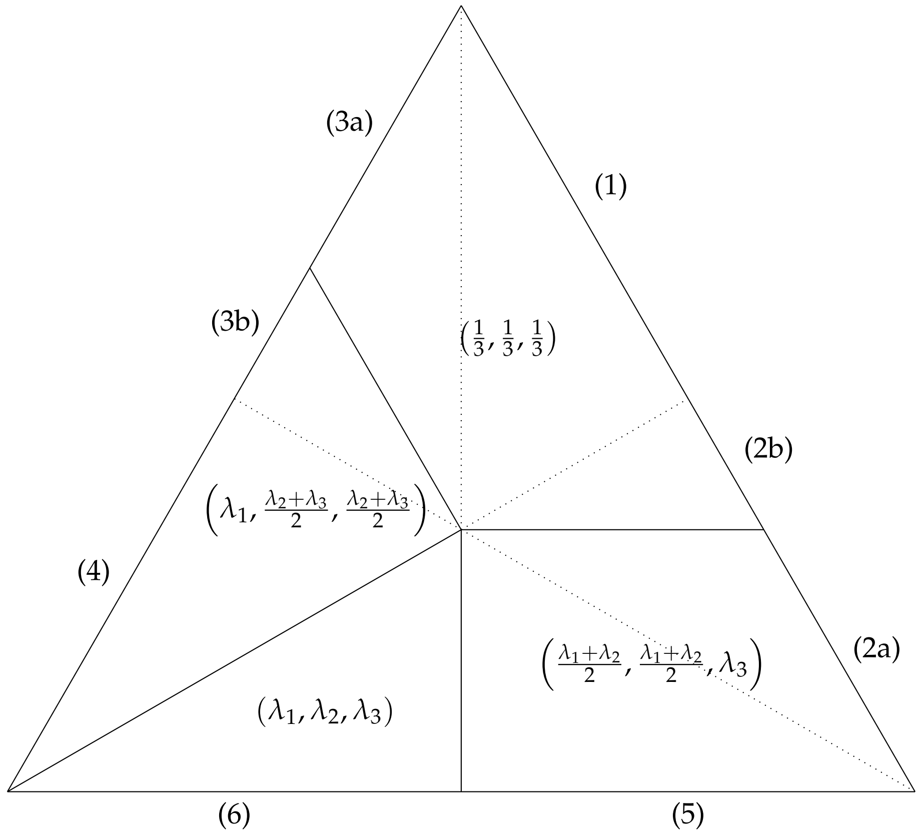

Figure 6: The barycentric plot of the generalised Hurwicz weights and the corresponding maximiser.

best case is necessary and sufficient for demanding the uniform distribution, whereas it is necessary but not sufficient in this case. We must also demand that $\lambda_1 + \lambda_2 \leq \frac{2}{3}$. So, for instance, if $\lambda_1 = \frac{1}{6}$, $\lambda_2 = \frac{2}{3}$, $\lambda_3 = \frac{1}{6}$, then GHC does not demand the uniform distribution. Rather, that assignment of weights belongs to (2a), and it demands $(\frac{\lambda_1+\lambda_2}{2}, \frac{\lambda_1+\lambda_2}{2}, \lambda_3) = (\frac{5}{12}, \frac{5}{12}, \frac{1}{6})$. In general, segments (2a) and (5) show that, if we continue to assign the majority of the weight to the worst and middle cases together, but shift it more towards the middle case, GHC no longer demands the uniform distribution, and rather demands that we pick a world and treat it as less likely than the others, which we treat equally. And segments (3b) and (4) show that, if we assign the majority of the weight to the best and worst cases together, but shift it more towards the best case, then GHC will demand that we pick a world and treat it as more likely than the other, which we treat equally (thus mirroring the demand of the original Hurwicz criterion). And (6) shows that, if we assign greatest weight to the best case, next greatest to the middle case, and least weight to the worst case, then GHC will simply demand that we adopt the weights as our credences.

So far, we have described what happens for the second-simplest case, namely, where there are just three worlds.[28] But what happens in general? This takes a little bit of terminology to describe. Given a sequence of numbers $A = (a_1, \ldots, a_n)$, let $\text{Av}(A) = \frac{a_1 + \ldots + a_n}{n}$. That is, $\text{Av}(A)$ is the arithmetic mean of the numbers in A. And given $1 \leq k \leq n$, let $A|_k = (a_1, \ldots, a_k)$. That is, $A|_k$ is the truncation of the sequence after a_k. Then we say that A *does not exceed its average* if, for each $1 \leq k \leq n$,

$$\text{Av}(A) \geq \text{Av}(A|_k)$$

That is, at no point in the sequence does the average of the numbers up to that point exceed the average of all the numbers in the sequence. We can now state our theorem:

Theorem 7 *Suppose $\Lambda = (\lambda_1, \ldots, \lambda_n)$ is a sequence of generalised Hurwicz weights. Then there is a sequence of subsequences $\Lambda_1, \ldots, \Lambda_m$ of Λ such that*

(i) $\Lambda = \Lambda_1 \frown \ldots \frown \Lambda_m$

(ii) $\text{Av}(\Lambda_1) \geq \ldots \geq \text{Av}(\Lambda_m)$

(iii) *each Λ_i does not exceed its average.*

Then,

$$(\underbrace{\text{Av}(\Lambda_1), \ldots, \text{Av}(\Lambda_1)}_{\text{length of } \Lambda_1}, \underbrace{\text{Av}(\Lambda_2), \ldots, \text{Av}(\Lambda_2)}_{\text{length of } \Lambda_2}, \ldots, \underbrace{\text{Av}(\Lambda_m), \ldots, \text{Av}(\Lambda_m)}_{\text{length of } \Lambda_m})$$

[28] In the simplest case, in which there are two worlds, GHC and Hurwicz's original criterion coincide.

maximises $H^\Lambda(\mathfrak{EU}(-))$ *among probabilistic credence functions* $C = (c_1, \ldots, c_n)$ *for which* $c_1 \geq \ldots \geq c_n$.

Here are a couple of examples:

- Suppose $\Lambda = \left(\frac{2}{10}, \frac{3}{10}, \frac{1}{10}, \frac{1}{10}, \frac{3}{10}\right)$.
  Then let $\Lambda_1 = \left(\frac{2}{10}, \frac{3}{10}\right)$, and $\Lambda_2 = \left(\frac{1}{10}, \frac{1}{10}, \frac{3}{10}\right)$. Then:
  (i) $\Lambda = \Lambda_1 \frown \Lambda_2$
  (ii) $\text{Av}(\Lambda_1) = \frac{1}{4} \geq \frac{1}{6} = \text{Av}(\Lambda_2)$
  (iii) Both Λ_1, Λ_2 do not exceed their averages.

  Thus, the maximiser in $\mathfrak{X}$ is $\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right)$. And thus the maximisers are precisely this and its permutations.

- Suppose $\Lambda = \left(\frac{5}{25}, \frac{7}{25}, \frac{5}{25}, \frac{3}{25}, \frac{5}{25}\right)$.
  Then let $\Lambda_1 = \left(\frac{5}{25}, \frac{7}{25}\right)$, $\Lambda_2 = \left(\frac{5}{25}\right)$, and $\Lambda_3 = \left(\frac{3}{25}, \frac{5}{25}\right)$. Then:
  (i) $\Lambda = \Lambda_1 \frown \Lambda_2 \frown \Lambda_3$
  (ii) $\text{Av}(\Lambda_1) = \frac{6}{25} \geq \text{Av}(\Lambda_2) = \frac{5}{25} \geq \text{Av}(\Lambda_3) = \frac{4}{25}$
  (iii) Λ_1, Λ_2, Λ_3 do not exceed their averages.

  Thus, the maximiser in $\mathfrak{X}$ is $\left(\frac{6}{25}, \frac{6}{25}, \frac{5}{25}, \frac{4}{25}, \frac{4}{25}\right)$. And thus the maximisers are precisely this and its permutations.

So we've seen which credence functions the GHC demands given different generalised Hurwicz weights. Let's now return to our taxonomy of permissivisms from Chapter 2 to see which one we can establish by appealing to GHC applied to measures of epistemic value that are generated by strictly proper scoring rules.

Of course, our version of permissivism concerns credences. Let's consider the interpersonal case. Whether we obtain permissivism, and if we do how radical it is, depends on which generalised Hurwicz weights it is permissible to have. For instance, in the case in which there are just three possible worlds, if it is only permissible to have weights $\lambda_1, \lambda_2, \lambda_3$ with $\lambda_1 \leq \frac{1}{3}$ and $\lambda_3 \geq \frac{2}{3}$, then we do not obtain permissivism at all. After all, any such weights demand the uniform distribution $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. But providing we allow something beyond that, we obtain permissivism. And the further beyond that we permit, the more radical the permissivism becomes. At its extreme, if we permit all possible generalised Hurwicz weights, we obtain the most radical possible permissivism—that is, for any probabilistic credence function, there are weights relative to which GHC permits it.

It's hard to know quite how to argue about which generalised Hurwicz weights are permissible. One way is to consider non-epistemic decisions, describe the preferences that the weights determine, and appeal to our judgment that they are rational preferences. Thus, for instance, consider the following two options, a and b:

	w_1	w_2	w_3
a_1	4	1	0
b_1	2	2	2

Then, if $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \frac{1}{3}$, $\lambda_2 = \frac{1}{6}$, then you'll prefer an option a to b. This doesn't seem beyond the pale. It is risk-inclined, for sure, but our Jamesian contention is that rationality does not prohibit that. We would not count someone irrational who risked getting 1 or even 0 utiles in order to open up the possibility of getting 4 utiles when the alternative can give them 2 utiles at most.

I take it that it is the very essence of William James' objection to Clifford that risk aversion is not the only rational attitude towards beliefs, just as it is not the only rational attitude towards other decisions in life. Absent reasons to the contrary, we should treat our choice of prior like any other decision formed in the absence of probabilities to guide us. This, we have said, requires that we use the Generalised Hurwicz Criterion. And that, in turn, says that we should pick generalised Hurwicz weights and then maximise the generalised Hurwicz score relative to them. As in the case of practical decisions, we permit at least some risk-inclined weights. And that is enough to give permissivism.

Nonetheless, perhaps there are limits. Here is one that we have already met: if any of our generalised Hurwicz weights is 0, then our preferences violate Strong Dominance. So I'm inclined to say that our weights must all be positive. And that entails that our prior credences similarly should all be positive—that is, at least at the beginning of our epistemic life, we should obey the Regularity Principle.

Here's a second: we might consider it irrational to prefer a_2 to b_2:

	w_1	w_2	w_3
a_2	$1,000,000$	0	0
b_2	$999,998$	$999,998$	$999,998$

But this is demanded if our Hurwicz weights are $(\lambda_1, \lambda_2, \lambda_3)$ and $\lambda_1 > \frac{999{,}998}{1{,}000{,}000}$. So perhaps there are limits to how risk-inclined it is rational to be. If there are, then there will be limits to the probability functions one might adopt as one's prior. For instance, if we prohibit $\lambda_1 > \frac{999{,}998}{1{,}000{,}000}$, then we rule out $\left(\frac{999{,}999}{1{,}000{,}000}, \frac{1}{2{,}000{,}000}, \frac{1}{2{,}000{,}000}\right)$ as a rational prior. It's hard to know how to adjudicate this question. On the one hand, I think we are confident that there are rational restrictions on the extremity of your risk attitudes—some seem simply beyond the pale. But on the other, there is no argument of which I am aware that pinpoints exactly what is so bad about such attitudes, nor where the pale is located. I'll leave this question open. If there are no such rational restrictions, we obtain the widest possible interpersonal permissivism compatible with Probabilism—namely, extreme subjective Bayesianism. If there are such rational restrictions, we obtain something slightly narrower. But either way, we still obtain permissivism.

Just as there might be limits to how risk-inclined it is rational to be, so there are likely limits to the level of rational risk-aversion. For instance, we might consider it irrational to prefer a_3 to b_3

	w_1	w_2	w_3
a_3	1	1	1
b_3	1,000,000	1,000,000	0

And yet we will prefer that if $\lambda_3 > \frac{999{,}999}{1{,}000{,}000}$. This latter point has less impact in the credal case, however, since even quite mild risk-aversion demands the uniform distribution. And so ruling out extreme risk aversion, which would also demand the uniform distribution, does not restrict the set of credence functions that may be justified using GHC.

Let's turn now to intrapersonal permissivism. That is, given your attitude to risk, encoded in your generalised Hurwicz weights, what is permissible for you? Well, this very much depends on what your weights are. Consider the regions in the simplex in Figure 6. If my weights are in (1), (3a), or (2b), there's no intrapersonal permissivism for me at all. Given my attitudes to risk, there's a unique credence function I must adopt and it's the uniform distribution. However, as I become less risk-averse by moving from the top right of the simplex into (2a) or (5), or into (3b) or (4), some intrapersonal permissivism opens up for me: for instance, in (2a) or (5), the following credence functions are permissible, all permutations of one another:

$$\left(\frac{\lambda_1+\lambda_2}{2}, \frac{\lambda_1+\lambda_2}{2}, \lambda_3\right), \left(\frac{\lambda_1+\lambda_2}{2}, \lambda_3, \frac{\lambda_1+\lambda_2}{2}\right), \left(\lambda_3 \frac{\lambda_1+\lambda_2}{2}, \frac{\lambda_1+\lambda_2}{2}\right)$$

And finally, if I become risk-inclined enough to occupy segment (6), there are six credence functions that are permissible:

$$\begin{array}{ccc} (\lambda_1, \lambda_2, \lambda_3) & (\lambda_2, \lambda_1, \lambda_3) & (\lambda_3, \lambda_1, \lambda_2) \\ (\lambda_1, \lambda_3, \lambda_2) & (\lambda_2, \lambda_3, \lambda_1) & (\lambda_3, \lambda_2, \lambda_1) \end{array}$$

In short, and very roughly speaking, the more risk-inclined and less risk-averse you are, the more is permissible for you.

The final part of our taxonomy asks how common permissivism is. How common are the bodies of evidence for which permissivism holds? For what proportion of bodies of evidence is the rational response to it not uniquely determined? But we can't answer that quite yet, because so far we have only spoken about how to pick priors. That is, we have considered how to respond to only one body of evidence, namely, the empty body that you have when you pick your priors. We will explore how to respond to other bodies of evidence in the next section.

5.3 Appendix: proofs

5.3.1 Proof of Theorem 7

Suppose $\mathfrak{EU}$ is a measure of epistemic value that is generated by the strictly proper scoring rule $\mathfrak{s}$. And suppose that Λ is the following sequence of generalised Hurwicz weights $0 \leq \lambda_1, \ldots, \lambda_n \leq 1$ with $\sum_{i=1}^{n} \lambda_i = 1$. In this section, we ask which are the credence functions C that maximise $H^\Lambda(\mathfrak{EU}(-))$.

We'll consider credence functions defined on the whole algebra $\mathcal{F}$ over the finite set of worlds $\mathcal{W}$. As we saw in Theorem 1 above, since GHC satisfies Weak Dominance, whatever maximises $H^\Lambda(\mathfrak{EU}(-))$ will be a probability function. Thus, $\sum_{w \in \mathcal{W}} C(w) = 1$ and $\sum_{w \in X} C(w) = C(X)$.

Next, recall that, if C maximises $H^\Lambda(\mathfrak{EU}(-))$, then so does any credence function obtained from C by a permutation π of the possible worlds in $\mathcal{W}$. If $\pi : \mathcal{W} \cong \mathcal{W}$, then define C_π such that $C_\pi(w) = C(\pi(w))$. Then, if C maximises $H^\Lambda(\mathfrak{EU}(-))$, so does C_π. From that, we know that there must be a maximiser C such that $c_1 \geq c_2 \geq \ldots \geq c_n$, where $c_i = C(w_i)$. It is this maximiser that we will show how to find.

Now, since $\mathfrak{EU}$ is generated by a strictly proper scoring rule, it is also truth-directed. That is, if $c_i > c_j$, then $\mathfrak{EU}(C, w_i) > \mathfrak{EU}(C, w_j)$. Thus, if $c_1 \geq c_2 \geq \ldots \geq c_n$, then

$$H^\Lambda(\mathfrak{EU}(C)) = \lambda_1 \mathfrak{EU}(C, w_1) + \ldots + \lambda_n \mathfrak{EU}(C, w_n)$$

This is what we seek to maximise. But notice that this is just the expectation of $\mathfrak{EU}(C)$ from the point of view of the probability function generated by $\Lambda = (\lambda_1, \ldots, \lambda_n)$. Throughout we will write $(p_1, \ldots, p_n)$ for the probability function that assigns p_i to world w_i. Thus, the probability function generated by $\Lambda = (\lambda_1, \ldots, \lambda_n)$ is the one that assigns λ_i to world w_i.

Now, Savage (1971) showed that, if $\mathfrak{s}$ is strictly proper and continuous, then there is a differentiable and strictly convex function φ such that, if P, Q are probabilistic credence functions, then

$$\begin{aligned} & \mathfrak{D}_\mathfrak{s}(P, Q) \\ := & \sum_{X \subseteq \mathcal{W}} \varphi(P(X)) - \sum_{X \subseteq \mathcal{W}} \varphi(Q(X)) - \sum_{X \subseteq \mathcal{W}} \varphi'(Q(X))(P(X) - Q(X)) \\ = & \sum_{i=1}^{n} p_i \mathfrak{EU}(P, w_i) - \sum_{i=1}^{n} p_i \mathfrak{EU}(Q, w_i) \end{aligned}$$

We call φ the *entropy function of* $\mathfrak{s}$.[29] So C maximises $H^\Lambda(\mathfrak{EU}(-))$ among credence functions C with $c_1 \geq \ldots \geq c_n$ iff C minimises $\mathfrak{D}_\mathfrak{s}(\Lambda, -)$ among

[29]For clear proof of this representation theorem, see (Predd et al., 2009).

credence functions C with $c_1 \geq \ldots \geq c_n$. We now use the KKT conditions to calculate which credence functions minimise $\mathfrak{D}_{\mathfrak{s}}(\Lambda, -)$ among credence functions C with $c_1 \geq \ldots \geq c_n$.

First, we define the following functions:

$$f(x_1, \ldots, x_n) = \mathfrak{D}_{\mathfrak{s}}((\lambda_1, \ldots, \lambda_n), (x_1, \ldots, x_n)) =$$

$$\sum_{X \subseteq \mathcal{W}} \varphi\left(\sum_{w_i \in X} \lambda_i\right) - \sum_{X \subseteq \mathcal{W}} \varphi\left(\sum_{w_i \in X} x_i\right) - \sum_{X \subseteq \mathcal{W}} \varphi'\left(\sum_{w_i \in X} x_i\right)\left(\sum_{w_i \in X} \lambda_i - \sum_{w_i \in X} x_i\right)$$

So the k^{th} entry in the vector ∇f is

$$\sum_{X \subseteq W : w_k \in X} \varphi''\left(\sum_{w_i \in X} x_i\right)\left(\sum_{w_i \in X} x_i - \sum_{w_i \in X} \lambda_i\right)$$

which we abbreviate $\theta_k(x)$.

Second, define the following functions:

$$\begin{array}{rcl}
h(x_1, \ldots, x_n) & = & x_1 + \ldots + x_n - 1 \\
g_1(x_1, \ldots, x_n) & = & x_2 - x_1 \\
g_2(x_1, \ldots, x_n) & = & x_3 - x_2 \\
\vdots & \vdots & \vdots \\
g_{n-2}(x_1, \ldots, x_n) & = & x_{n-1} - x_{n-2} \\
g_{n-1}(x_1, \ldots, x_n) & = & x_n - x_{n-1}
\end{array}$$

Then

$$\begin{array}{rcl}
\nabla h & = & \langle 1, 1, 1, 1 \ldots, 1, 1, 1, 1 \rangle \\
\nabla g_1 & = & \langle -1, 1, 0, 0 \ldots, 0, 0, 0, 0 \rangle \\
\nabla g_2 & = & \langle 0, -1, 1, 0, \ldots, 0, 0, 0, 0 \rangle \\
\nabla g_3 & = & \langle 0, 0, -1, 1, \ldots, 0, 0, 0, 0 \rangle \\
\vdots & \vdots & \vdots \\
\nabla g_{n-3} & = & \langle 0, 0, 0, 0, \ldots, -1, 1, 0, 0 \rangle \\
\nabla g_{n-2} & = & \langle 0, 0, 0, 0, \ldots, 0, -1, 1, 0 \rangle \\
\nabla g_{n-1} & = & \langle 0, 0, 0, 0, \ldots, 0, 0, -1, 1 \rangle
\end{array}$$

So the KKT theorem says that $x_1, \ldots, x_n$ is a maximiser iff there are $\lambda, \mu_1, \ldots, \mu_{n-1}$ with $0 \leq \mu_1, \ldots, \mu_{n-1}$ such that

$$\nabla f(x_1, \ldots, x_n) + \sum_{i=1}^{n-1} \mu_i \nabla g_i(x_1, \ldots, x_n) + \lambda \nabla h(x_1, \ldots, x_n) = 0$$

That is, iff there are λ and $0 \leq \mu_1, \ldots, \mu_{n-1}$ such that

$$\begin{array}{rcl} \theta_1(x) - \mu_1 + \lambda & = & 0 \\ \theta_2(x) - \mu_2 + \mu_1 + \lambda & = & 0 \\ \vdots & \vdots & \vdots \\ \theta_{n-1}(x) - \mu_{n-1} + \mu_{n-2} + \lambda & = & 0 \\ \theta_n(x) + \mu_{n-1} + \lambda & = & 0 \end{array}$$

By summing the identities above, we get

$$\lambda = -\frac{1}{n} \sum_{j=1}^{n} \theta_j(x)$$

and

$$\mu_k = \sum_{i=1}^{k} \theta_i(x) - \frac{k}{n} \sum_{i=1}^{n} \theta_i(x)$$

Now, suppose that there is a sequence of subsequences $\Lambda_1, \ldots, \Lambda_m$ of Λ such that

(i) $\Lambda = \Lambda_1 \frown \ldots \frown \Lambda_m$

(ii) $\text{Av}(\Lambda_1) \geq \ldots \geq \text{Av}(\Lambda_m)$

(iii) each Λ_j does not exceed its average.

And let

$$P = (\underbrace{\text{Av}(\Lambda_1), \ldots, \text{Av}(\Lambda_1)}_{\text{length of } \Lambda_1}, \underbrace{\text{Av}(\Lambda_2), \ldots, \text{Av}(\Lambda_2)}_{\text{length of } \Lambda_2}, \ldots, \underbrace{\text{Av}(\Lambda_m), \ldots, \text{Av}(\Lambda_m)}_{\text{length of } \Lambda_m})$$

That is, given $1 \leq i \leq n$, if λ_i is in Λ_j, then $p_i = \text{Av}_j$, where we write Av_j to abbreviate $\text{Av}(\Lambda_j)$.

We will now show that, if we let $x_i = p_i$, for each $1 \leq i \leq n$, and we take $\lambda, \mu_1, \ldots, \mu_{n-1}$ to be defined as above, then

(i) $\lambda = 0$; and

(ii) $\mu_i \geq 0$.

Thanks to the KKT theorem, this is sufficient to complete the theorem.

Our proof depends on an identity, which we will derive now. In the derivation, we make use of a particular way of categorising propositions. We say that the type of a proposition $X \subseteq W$ is the sequence $\beta = (\beta_1, \ldots, \beta_m)$ such that $|X \cap \Lambda_j| = \beta_j$, for $1 \leq j \leq m$. Given a proposition $X \subseteq W$ and

$\beta = (\beta_1, \ldots, \beta_m)$, with $\beta_j \leq \Lambda_j$ for each $1 \leq j \leq m$, we write $X \in \beta$ iff the type of X is β. Notice that, if X is in β, then

$$\sum_{w_i \in X} p_i = \sum_{j=1}^{m} \beta_j \mathrm{Av}_j$$

Then, for $1 \leq j^* \leq m$ and $1 \leq k^* \leq n$ with λ_{k^*} in Λ_{l^*},

$$\begin{aligned}
\sum_{k \in \Lambda_{j^*}|_{k^*}} \theta_k(p) &= \sum_{k \in \Lambda_{j^*}|_{k^*}} \sum_{\substack{X \subseteq W \\ w_i \in X}} \varphi'' \left(\sum_{w_j \in X} p_j \right) \left(\sum_{w_j \in X} p_j - \sum_{w_j \in X} \lambda_j \right) \\
&= \sum_{\beta} \sum_{k \in \Lambda_{j^*}|_{k^*}} \sum_{\substack{X \in \beta \\ w_k \in X}} \varphi'' \left(\sum_{w_i \in X} p_i \right) \left(\sum_{w_i \in X} p_i - \sum_{w_i \in X} \lambda_i \right) \\
&= \sum_{\beta} \sum_{k \in \Lambda_{j^*}|_{k^*}} \sum_{\substack{X \in \beta \\ w_k \in X}} \varphi'' \left(\sum_{j=1}^{m} \beta_j \mathrm{Av}_j \right) \left(\sum_{j=1}^{m} \beta_j \mathrm{Av}_j - \sum_{w_i \in X} \lambda_i \right) \\
&= \sum_{\beta} \varphi'' \left(\sum_{j=1}^{m} \beta_j \mathrm{Av}_j \right) \sum_{k \in \Lambda_{j^*}|_{k^*}} \sum_{\substack{X \in \beta \\ w_k \in X}} \left(\sum_{j=1}^{m} \beta_j \mathrm{Av}_j - \sum_{w_i \in X} \lambda_i \right)
\end{aligned}$$

Now, for each λ_k in $\Lambda_{j^*}|_{k^*}$, the number of propositions X in β with $w_k \in X$ is

$$\prod \binom{\Lambda_{j^*} - 1}{\beta_{j^*} - 1} = \prod \binom{\Lambda_{j^*}}{\beta_{j^*}} \frac{\beta_{j^*}}{\Lambda_{j^*}}$$

where we abuse notation and write Λ_j for $|\Lambda_j|$, that is, the length of sequence Λ_j. Note, this number doesn't depend on k. So the number summands in $\sum_{k \in \Lambda_{j^*}|_{k^*}} \sum_{\substack{X \in \beta \\ w_k \in X}}$ is

$$\Lambda_{j^*}|_{k^*} \prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_{j^*}}{\Lambda_{j^*}}$$

Thus,

$$\begin{aligned}
& \sum_{k \in \Lambda_{j^*}|_{k^*}} \sum_{\substack{X \in \beta \\ w_k \in X}} \sum_{j=1}^{m} \beta_j \mathrm{Av}_j \\
= \; & \Lambda_{j^*}|_{k^*} \prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_{j^*}}{\Lambda_{j^*}} \sum_{j=1}^{m} \beta_j \mathrm{Av}_j \\
= \; & \prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_{j^*}}{\Lambda_{j^*}} \sum_{j=1}^{m} \beta_j \Lambda_{j^*}|_{k^*} \mathrm{Av}_j
\end{aligned}$$

Next, consider

$$\sum_{i \in \Lambda_l|k} \sum_{\substack{X \in \beta \\ w_i \in X}} \sum_{w_j \in X} \lambda_j$$

and ask, for each $1 \le i \le n$, how many times λ_i occurs in this summation. Here's the answer:

- If λ_i is in Λ_j for $j \neq j^*$, then λ_i occurs this many times:

$$\sum_{\substack{j=1 \\ j \neq j^*}}^{m} \prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_j}{\Lambda_j} \frac{\beta_{j^*}}{\Lambda_{j^*}} \Lambda_{j^*}|_{k^*}$$

- If λ_i is in $\Lambda_{j^*} - \Lambda_{j^*}|_{k^*}$, then λ_i occurs this many times:

$$\prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_{j^*}(\beta_{j^*}-1)}{\Lambda_{j^*}(\Lambda_{j^*}-1)} \Lambda_{j^*}|_{k^*}$$

- If λ_i is in $\Lambda_{j^*}|_{k^*}$, then λ_i occurs this many times:

$$\prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_{j^*}}{\Lambda_{j^*}} \left(1 + \frac{\beta_{j^*}-1}{\Lambda_{j^*}-1}(\lambda_{j^*}|_{k^*} - 1)\right)$$

So

$$\begin{aligned}
& \sum_{k \in \Lambda_{j^*}|_{k^*}} \sum_{\substack{X \in \beta \\ w_k \in X}} \sum_{w_i \in X} \lambda_i \\
= \; & \sum_{\substack{j=1 \\ j \neq j^*}}^{m} \prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_j}{\Lambda_j} \frac{\beta_{j^*}}{\Lambda_{j^*}} \Lambda_{j^*}|_{k^*} \sum_{\lambda_i \in \Lambda_j} \lambda_i + \\
& \prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_{j^*}(\beta_{j^*}-1)}{\Lambda_{j^*}(\Lambda_{j^*}-1)} \Lambda_{j^*}|_{k^*} \sum_{\lambda_i \in \Lambda_{j^*} - \Lambda_{j^*}|_{k^*}} \lambda_i + \\
& \prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_{j^*}}{\Lambda_{j^*}} \left(1 + \frac{\beta_{j^*}-1}{\Lambda_{j^*}-1}(\lambda_{j^*}|_{k^*} - 1)\right) \sum_{\lambda_i \in \Lambda_{j^*}|_{k^*}} \lambda_i \\
= \; & \prod_{j=1}^{m} \binom{\Lambda_j}{\beta_j} \frac{\beta_{j^*}}{\Lambda_{j^*}} \left(\sum_{\substack{j=1 \\ j \neq j^*}}^{m} \frac{\beta_j}{\Lambda_j} \Lambda_{j^*}|_{k^*} \sum_{\lambda_i \in \Lambda_j} \lambda_i + \right. \\
& \left. \frac{\beta_{j^*}-1}{\Lambda_{j^*}-1} \Lambda_{j^*}|_{k^*} \sum_{\lambda_i \in \Lambda_{j^*} - \Lambda_{j^*}|_{k^*}} \lambda_i + \left(1 + \frac{\beta_{j^*}-1}{\Lambda_{j^*}-1}(\lambda_{j^*}|_{k^*} - 1)\right) \sum_{\lambda_i \in \Lambda_{j^*}|_{k^*}} \lambda_i \right)
\end{aligned}$$

Now, if $j \neq j^*$, then

$$\Lambda_{j^*}|_{k^*}\beta_j \text{Av}_j - \frac{\beta_j}{\Lambda_j}\Lambda_{j^*}|_{k^*}\sum_{i\in\Lambda_j}\lambda_i = \Lambda_{j^*}|_{k^*}\beta_j\text{Av}_j - \Lambda_{j^*}|_{k^*}\beta_j\text{Av}_j = 0$$

While

$$\begin{aligned}
& \Lambda_{j^*}|_{k^*}\beta_{j^*}\text{Av}_{j^*} - \left(\frac{\beta_{j^*}-1}{\Lambda_{j^*}-1}\Lambda_{j^*}|_{k^*}\sum_{\lambda_i\in\Lambda_{j^*}-\Lambda_{j^*}|_{k^*}}\lambda_i\right) - \\
& \left(1+\frac{\beta_{j^*}-1}{\Lambda_{j^*}-1}(\Lambda_{j^*}|_{k^*}-1)\right)\sum_{\lambda_i\in\Lambda_{j^*}|_{k^*}}\lambda_i \\
= \; & \Lambda_{j^*}|_{k^*}\beta_{j^*}\text{Av}_{j^*} - \left(\frac{\beta_{j^*}-1}{\Lambda_{j^*}-1}\Lambda_{j^*}|_{k^*}\sum_{\lambda_i\in\Lambda_{j^*}}\lambda_i\right) - \left(1-\frac{\beta_{j^*}-1}{\Lambda_{j^*}-1}\right)\sum_{\lambda_i\in\Lambda_{j^*}|_{k^*}}\lambda_i \\
= \; & \Lambda_{j^*}|_{k^*}\text{Av}_{j^*}\left(\beta_{j^*}-\frac{\Lambda_{j^*}(\beta_{j^*}-1)}{\Lambda_{j^*}-1}\right) - \frac{\Lambda_{j^*}-\beta_{j^*}}{\Lambda_{j^*}-1}\sum_{\lambda_i\in\Lambda_{j^*}|_{k^*}}\lambda_i \\
= \; & \frac{\Lambda_{j^*}-\beta_{j^*}}{\Lambda_{j^*}-1}\left(\Lambda_{j^*}|_{k^*}\text{Av}_{j^*} - \sum_{\lambda_i\in\Lambda_{j^*}|_{k^*}}\lambda_i\right)
\end{aligned}$$

Putting all of this together, we have: for $1 \leq j^* \leq m$ and $1 \leq k^* \leq n$ with λ_{k^*} in Λ_{l^*},

$$\sum_{k\in\Lambda_{j^*}|_{k^*}}\theta_k(p) = \frac{\Lambda_{j^*}-\beta_{j^*}}{\Lambda_{j^*}-1}\left(\Lambda_{j^*}|_{k^*}\text{Av}_{j^*} - \sum_{\lambda_i\in\Lambda_{j^*}|_{k^*}}\lambda_i\right)$$

So, in particular,

$$\sum_{k\in\Lambda_{j^*}}\theta_k(p) = \frac{\Lambda_{j^*}-\beta_{j^*}}{\Lambda_{j^*}-1}\left(\Lambda_{j^*}\text{Av}_{j^*} - \sum_{\lambda_i\in\Lambda_{j^*}}\lambda_i\right) = 0$$

And so:

$$\lambda = -\frac{1}{n}\sum_{k=1}^{n}\theta_k(p) = 0$$

And, if $1 \leq j \leq m$, $1 \leq k \leq n$, and λ_k is in Λ_j

$$\mu_k = \sum_{i=1}^{k}\theta_i(p) - \frac{k}{n}\sum_{i=1}^{n}\theta_i(x) = \frac{\Lambda_j-\beta_j}{\Lambda_j-1}\left(\Lambda_j|_k\text{Av}_j - \sum_{\lambda_i\in\Lambda_j|_k}\lambda_i\right)$$

which is non-negative iff $|\Lambda_j|_k|\text{Av}_l - \sum_{\lambda_i\in\Lambda_j|_k}\lambda_i \geq 0$ iff $\text{Av}(\Lambda_j) \geq \text{Av}(\Lambda_j|_k)$. But, by assumption, this is true, since each Λ_j does not exceed its average. So, in conclusion, P maximises $H^{\Lambda}(\mathfrak{EU}(-))$, as required.

This completes our proof that, if there is a sequence of subsequences of Λ that satisfy (i), (ii), (iii) from above, then they provide the maximiser. We now show that such subsequences always exist. We proceed by induction.

BASE CASE $n = 1$. Then it is clearly true with the subsequence $\Lambda_1 = \Lambda$.

INDUCTIVE STEP Suppose it is true for all sequences $\Lambda = (\lambda_1, \dots, \lambda_n)$ of length n. Now consider a sequence $(\lambda_1, \dots, \lambda_n, \lambda_{n+1})$. Then, by the inductive hypothesis, there is a sequence of sequences $\Lambda_1, \dots, \Lambda_m$ such that

(i) $\Lambda \frown (\lambda_{n+1}) = \Lambda_1 \frown \dots \frown \Lambda_m \frown (\lambda_{n+1})$

(ii) $\mathrm{Av}(\Lambda_1) \geq \dots \geq \mathrm{Av}(\Lambda_m)$

(iii) each Λ_j does not exceed its average.

Now, first, suppose $\mathrm{Av}(\Lambda_m) \geq \lambda_{n+1}$. Then let $\Lambda_{m+1} = (\lambda_{n+1})$ and we're done.

So, second, suppose $\mathrm{Av}(\Lambda_m) < \lambda_{n+1}$. Then we find the greatest k such that

$$\mathrm{Av}(\Lambda_k) \geq \mathrm{Av}(\Lambda_{k+1} \frown \dots \frown \Lambda_m \frown (\lambda_{n+1}))$$

Then we let $\Lambda^*_{k+1} = \Lambda_{k+1} \frown \dots \frown \Lambda_m \frown (\lambda_{n+1})$. Then we can show that

(i) $(\lambda_1, \dots, \lambda_n, \lambda_{n+1}) = \Lambda_1 \frown \Lambda_2 \frown \dots \frown \Lambda_k \frown \Lambda^*_{k+1}$.

(ii) Each $\Lambda_1, \dots, \Lambda_k, \Lambda^*_{k+1}$ does not exceed average.

(iii) $\mathrm{Av}(\Lambda_1) \geq \mathrm{Av}(\Lambda_2) \geq \dots \geq \mathrm{Av}(\Lambda_k) \geq \mathrm{Av}(\Lambda^*_{k+1})$.

(i) and (iii) are obvious. So we prove (ii). In particular, we show that Λ^*_{k+1} does not exceed average. We assume that each subsequence Λ_j starts with Λ_{i_j+1}

- Suppose $i \in \Lambda_{k+1}$. Then, since Λ_{k+1} does not exceed average,

$$\mathrm{Av}(\Lambda_{k+1}) \geq \mathrm{Av}(\Lambda_{k+1}|_i)$$

But, since k is the greatest number such that

$$\mathrm{Av}(\Lambda_k) \geq \mathrm{Av}(\Lambda_{k+1} \frown \dots \frown \Lambda_m \frown (\lambda_{n+1}))$$

We know that

$$\mathrm{Av}(\Lambda_{k+2} \frown \dots \frown \Lambda_m \frown (\lambda_{n+1})) > \mathrm{Av}(\Lambda_{k+1})$$

So

$$\mathrm{Av}(\Lambda_{k+1} \frown \dots \frown \Lambda_m \frown (\lambda_{n+1})) > \mathrm{Av}(\Lambda_{k+1})$$

So

$$\mathrm{Av}(\Lambda_{k+1} \frown \dots \frown \Lambda_m \frown (\lambda_{n+1})) > \mathrm{Av}(\Lambda_{k+1}|_i)$$

- Suppose $i \in \Lambda_{k+2}$. Then, since Λ_{k+2} does not exceed average,

  $$\text{Av}(\Lambda_{k+2}) \geq \text{Av}(\Lambda_{k+2}|_i)$$

  But, since k is the greatest number such that

  $$\text{Av}(\Lambda_k) \geq \text{Av}(\Lambda_{k+1} \frown \ldots \frown \Lambda_m \frown (\lambda_{n+1}))$$

  We know that

  $$\text{Av}(\Lambda_{k+3} \frown \ldots \frown \Lambda_m \frown (\lambda_{n+1})) > \text{Av}(\Lambda_{k+2})$$

  So

  $$\text{Av}(\Lambda_{k+1} \frown \ldots \frown \Lambda_m \frown (\lambda_{n+1})) > \text{Av}(\Lambda_{k+2}|_i)$$

  But also, from above,

  $$\text{Av}(\Lambda_{k+1} \frown \ldots \frown \Lambda_m \frown (\lambda_{n+1})) > \text{Av}(\Lambda_{k+1})$$

  So

  $$\text{Av}(\Lambda_{k+1} \frown \ldots \frown \Lambda_m \frown (\lambda_{n+1})) > \text{Av}(\Lambda_{k+1} \frown \Lambda_{k+2}|_i)$$

- And so on.

This completes the proof. □

In brief...

In previous chapters, we have argued that, when we pick our priors, we should do so using (i) strictly proper epistemic utility functions to measure the epistemic value of a credal state, and (ii) a decision rule from the family of risk-sensitive decision rules called the Generalised Hurwicz Criterion. Recall: each member of this family is picked out by a sequence of generalised Hurwicz weights, and these encode our attitudes to risk. In this chapter, we derive the consequences of picking our priors in this way. Some of the upshots:

- For many Hurwicz weights that encode attitudes to risk that range from risk-averse to risk-neutral, if you apply GHC with these weights to your choice of priors, it demands that you adopt the uniform prior—that is, the one that divides your credences equally over all worlds.

- As we move from these risk-averse and risk-neutral weights to more risk-inclined ones, the resulting versions of GHC demand priors that are not uniform—they permit priors that assign more credence to some worlds than to others; and as the priors becoming more and more risk-inclined, they permit priors that assign a lot of their credence to some worlds, and much less to others.

- In the end, for any probabilistic prior credence function, there is a set of generalised Hurwicz weights—and so, a set of attitudes to epistemic risk—for which that credence function is among those rationally permitted by GHC for an individual with those weights.

  This gives a wide version of interpersonal permissivism about rational prior credences.

- If your weights demand the uniform distribution then for you rationality is intrapersonally impermissive concerning prior credences—there is a single rational prior for you, and it is the uniform distribution.

- However, if your weights do not demand the uniform distribution, then for you rationality is intrapersonally permissive about priors—after all, if one credence function is permitted by GHC, so is any permutation of it; that is, if one prior assignment of credences to possible worlds is permitted so is any assignment that assigns the same credences but to different worlds.

  This gives a narrower, but sometimes still quite wide version of intrapersonal permissivism about rational prior credences.

6 Epistemic risk and picking posteriors

So, having set our priors using GHC, how should we set our posteriors? That is, how should we set our credences at times other than the very beginning of our epistemic life, when our body of total evidence might no longer be empty? In this chapter, we set out a variety of possible approaches to this question. In the end, I'll opt for one that builds on insights from work by Kenny Easwaran, as well as ideas that Hannes Leitgeb and I developed ten years ago, and which Dmitri Gallow has recently improved (Leitgeb & Pettigrew, 2010b; Easwaran, 2014; Gallow, 2019). But it is interesting to work through the other possibilities to see why they are problematic. Those impatient for the truth and indifferent to interesting falsehoods can skip straight to Section 6.3.

6.1 GHC forever

One natural proposal is simply to apply GHC at each point in your epistemic life, not just at the initial point. There is a wrinkle here, however. Let's say that $\mathcal{W} = \{w_1, \ldots, w_n\}$ is the set of possible worlds about which you have an opinion. As you obtain new evidence, you rule out those worlds that are incompatible with your evidence. When you apply GHC at the initial time in your epistemic life, you assign n Hurwicz weights, $\lambda_1, \ldots, \lambda_n$, one for each place in the ordering of worlds from best to worst, and you apply GHC with those weights. Now, suppose you learn $E = \{w_1, \ldots, w_m\}$, for some $m < n$, and you wish to apply GHC again. How do you do this?

On the one hand, if you simply use the same n Hurwicz weights, GHC will permit the same credence functions that it permitted at the initial time. If only the uniform distribution was permissible then, only the uniform distribution will be permissible now; and so on. But then it makes no difference that you have learned E. What it is rational for you to believe will be sensitive neither to whether you have received any new evidence nor to what evidence you have received if you have. What's more, among the permissible posteriors there will typically be none that are obtained from your prior by applying Bayes' rule to your evidence; and indeed none even that make you certain of the evidence that you have learned. If I am required to have the uniform distribution as my prior, and then required again to have it as my posterior after learning E, then my posterior does not evolve from my prior by Bayes' Rule, and it does not even assign certainty to E—it assigns credence $\frac{m}{n} < 1$.

On the other hand, you can't apply GHC to the new, more restricted set of epistemically possible worlds, since there's no principled way to extract your m Hurwicz weightings $\lambda'_1, \ldots, \lambda'_m$ relative to the set of possible worlds $E = \{w_1, \ldots, w_m\}$ from your old weightings relative to the set

$\mathcal{W} = \{w_1, \ldots, w_n\}$. Suppose $\mathcal{W} = \{w_1, w_2, w_3\}$ and $E = \{w_1, w_2\}$. How should we obtain λ'_1, λ'_2 from $\lambda_1, \lambda_2, \lambda_3$? Should we sum λ_1 and λ_2, our original generalised Hurwicz weights for the best and second-best case, to give λ'_1, our new generalised Hurwicz weights for the best case? Or should we let λ_1 be λ'_1? Or something else?

I think the best proposal in this vein is to apply GHC with the original n weights, while restricting the set of available posteriors to those that respect your evidence by assigning it maximal credence. That is, we must pick a posterior that maximises your generalised Hurwicz score $H^{\Lambda}(\mathfrak{EU}(-))$ relative to your weights *among the credence functions C with $C(E) = 1$*. As we'll see, however, this is often incompatible with Bayes' Rule.

For the sake of providing concrete examples, we'll assume:

- the set of possible worlds is $\mathcal{W} = \{w_1, w_2, w_3\}$;
- your total evidence at the later time is $E = \{w_1, w_2\}$;
- your generalised Hurwicz weights are $\lambda_1, \lambda_2, \lambda_3$.

There are three cases I'd like to consider. They don't exhaust all the options, but they nicely illustrate the possibilities and problems with this approach. They are distinguished by the Hurwicz weights that you assign

(i) $\lambda_1 \leq \lambda_2$ and $\frac{1}{3} \leq \lambda_3$—that is, your Hurwicz weights lie in region (1) or (2b) from Figure 6 above;

(ii) $\lambda_1 \leq \lambda_2$ and $\lambda_3 < \frac{1}{3}$—that is, they lie in region (2a) or (5) from Figure 6 above;

(iii) $\lambda_2 < \lambda_1$ and $\lambda_1 < \frac{1}{3}$—that is, they lie region (3a) from Figure 6 above.

Let's treat them in turn.

Suppose (i) $\lambda_1 \leq \lambda_2$ and $\frac{1}{3} \leq \lambda_3$. Then, however you measure epistemic utility, applying GHC without any restrictions requires you to adopt prior $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. And, again regardless of how you measure epistemic utility, applying GHC with the restriction $c(E) = 1$ requires you to adopt $(\frac{1}{2}, \frac{1}{2}, 0)$.[30] And the latter is the result of applying Bayes' Rule to the former. What's more, this will work for whatever evidence you acquire. If you start with $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, then, for any evidence you might acquire, you will get the same posterior whether you update your prior on this evidence using Bayes' Rule or whether you apply GHC with the restriction that your posterior must be certain of this evidence.

Next, suppose (ii) $\lambda_1 \leq \lambda_2$ and $\lambda_3 < \frac{1}{3}$. Then GHC requires you to adopt $(\frac{\lambda_1+\lambda_2}{2}, \frac{\lambda_1+\lambda_2}{2}, \lambda_3)$ or one of its permutations as your prior. And, again, applying GHC with the restriction $c(E) = 1$ requires you to adopt $(\frac{1}{2}, \frac{1}{2}, 0)$.[31]

[30] I state and prove the general result as Theorem 9 in Section 6.5 below.

[31] Again, this is proved within Theorem 9 in Section 6.5 below.

And this is the result of applying Bayes' Rule to $(\frac{\lambda_1+\lambda_2}{2}, \frac{\lambda_1+\lambda_2}{2}, \lambda_3)$. However, it is not the result of applying Bayes' Rule to $(\frac{\lambda_1+\lambda_2}{2}, \lambda_3, \frac{\lambda_1+\lambda_2}{2})$, which GHC also permits as a prior with these generalised Hurwicz weights. So we cannot say anything as strong as we did for case (i). Bayes' Rule and GHC will not always coincide. Nonetheless, we can say this: for someone with these generalised Hurwicz weights, there is always a route through their epistemic life that satisfies GHC at each stage and evolves by updating on evidence using Bayes' Rule at each stage. That is, for any such weights and any future body of evidence, there are priors permitted by GHC, and posteriors permitted by the version of GHC that restricts to posteriors that are certain of the evidence, such that the posteriors are obtained from the priors by applying Bayes' Rule to the evidence.

Finally, suppose (iii) $\lambda_2 < \lambda_1$ and $\frac{2}{3} < \lambda_3$. Then GHC requires your prior to be $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, and applying Bayes' Rule to that when you learn E would give $(\frac{1}{2}, \frac{1}{2}, 0)$. However, there are no strictly proper scoring rules for which applying GHC with those weights and the restriction $c(E) = 1$ gives $(\frac{1}{2}, \frac{1}{2}, 0)$.[32] So, for these generalised Hurwicz weights there can be no route through your epistemic life that satisfies both GHC and Bayes' Rule.

Of course, you might ask why we should be so concerned to update in the way that Bayes' Rule recommends at each stage. One reason comes from the famous diachronic sure loss argument due to David Lewis (1999). Like other sure loss arguments for credal principles, such as Ramsey's and de Finetti's famous sure loss argument for Probabilism, Lewis' argument relies on the following claim about the bets that your credences require you to accept: if you have credence $0 \leq r \leq 1$ in proposition X, you are rationally required to pay any amount less than rS utiles for a bet that pays out S utiles if X is true and 0 utiles if X is false. Lewis then shows that, if you plan to update in some way other than by Bayes' Rule, there are bets that your priors require you to accept and bets that your planned posteriors will require you to accept that, taken together, will lose you utility for sure. I have doubts about this argument. It tells you that, were you to face certain choices, your priors and your posteriors would require you to choose in a way that is guaranteed to serve you badly. But it says nothing about how they will require you to choose when you face other choices. Perhaps your priors and planned posteriors will serve you very well in those situations. The problem is that the diachronic sure loss argument does not tell us whether there is a trade-off here that might make it reasonable to keep the prior credence function and updating plan even though they are, taken together, vulnerable to the diachronic sure loss. Perhaps there is some way in which the prior and plan compensate for their poor performance in the face of the specific pair of decision problems that underpin the sure loss argument. Perhaps they perform extremely well when presented with an-

[32] Another consequence of Theorem 9 in Section 6.5 below.

other pair of decision problems, and perhaps we are more likely to face this alternative pair. I've explored this question in more detail elsewhere, so I'll leave it here, particularly since these pragmatic arguments for credal norms are not our main concern here.[33]

Instead, let's turn to one of the epistemic utility arguments for planning to update by Bayes' Rule that we met in Section 3.3. According to the result that Ray Briggs and I presented, if you plan to update your prior in any way other than by Bayes' Rule, there will be an alternative prior and an alternative updating plan that, taken together, have greater total epistemic utility than your prior and plan, taken together. What's more, if you always plan to update by Bayes' Rule, this won't happen.

A notable feature of this argument is that, if it works at all, it establishes a wide scope version of the norm that says you should update by Bayes' Rule. That is, it does not establish: if your prior is P, then you should plan to pick your posteriors by conditioning P on your evidence. Rather, it establishes: you ought not to have prior P and plan to pick your posteriors by something other than conditioning P on your evidence. But of course this does tell against the person who plans to use GHC at each point of time, simply restricting at each point to those credence functions that are certain of your total evidence at that time. At least if they are in situation (iii) from above, the argument above shows that the prior and updating plan that GHC requires them to have are, taken together, dominated.

This argument shows that, in the epistemic context, planning to make repeated use of GHC with certain Hurwicz weights leads to dominated choices. So what gives? The repeated use of GHC? Or the troublesome Hurwicz weights? We might retain GHC as our decision rule at each stage of our epistemic life and badge as rationally impermissible the attitudes to risk that are encoded in the Hurwicz weights that lead to dominated choices—e.g. those in (iii) above. Or we might say that GHC applies at only one time, namely, the beginning of your epistemic life, when you use it to pick your priors. I opt to retain for the latter for two reasons.

First, as we saw in our discussion of Strong Summation above, we have good reason *not* to apply GHC more than once; and it is a reason that is quite independent of the result that Briggs and I offer. After all, as we saw in footnote 25, there is only one set of generalised Hurwicz weights that avoids the problem of diachronic incoherence: it is the uniform weights, $\lambda_1 = \lambda_2 = \ldots = \lambda_n = \frac{1}{n}$. For any others, there will be sequences of choices in the face of which GHC will require you to choose a dominated sequence of options. So, again, if we wish to retain our Jamesian permissivism about attitudes to risk, we should restrict our use of GHC to a single time.

[33]I discuss this objection to the diachronic sure loss argument in (Pettigrew, 2020). And I discuss a related objection to exploitation arguments for rationality more generally in (Pettigrew, 2019, Chapter 13).

Second, as we have noted before, using GHC only at the beginning of your epistemic life makes sense. Once we have used GHC to choose our priors, we are in a situation in which we have probabilistic opinions that we should use to make future decisions. Indeed, part of our purpose in picking priors is to do exactly this with them.

6.2 Priors and plans together

So, applying GHC at each point in time leads us in the epistemic case—as in the practical case—to the possibility of dominated decisions, unless our Hurwicz weights lie in a particular risk-averse region of the simplex. But the argument that establishes this might give us an idea. In that argument, we considered the combination of an individual's prior credence function and her updating plan. And we applied the decision-theoretic principle of Weak Dominance to show that, if you plan to obtain your posterior credence function by updating your prior credence function in some way other than by Bayes' Rule, then you are irrational, because there is an alternative prior and an alternative plan that are jointly more valuable than yours, epistemically speaking, at every world. Perhaps, then, as well as applying the Weak Dominance principle when we choose these prior-plan pairs, we might apply the stronger principle GHC as well. As we will see, the problem with this approach is that the priors it requires you to have are determined not only by your generalised Hurwicz weights, but also by the partition from which your evidence will come. And this is not something you typically know at the point when you are choosing your priors.

Let's see how this works in a specific case. As usual, we assume there are just three worlds: w_1, w_2, w_3. And we suppose that your Hurwicz weights are $\lambda_1 = \frac{1}{6}$, $\lambda_2 = \frac{1}{2}$, and $\lambda_3 = \frac{1}{3}$. And now consider two partitions of possible evidence:

(i) $\mathcal{E} = \{E_1 = \{w_1\}, E_2 = \{w_2\}, E_3 = \{w_3\}\}$;

(ii) $\mathcal{E}' = \{E'_1 = \{w_1, w_2\}, E'_2 = \{w_3\}\}$.

Now, we know by the dominance argument that Briggs and I offered that GHC requires you to plan to update by Bayes' Rule—since dominated options never maximise generalised Hurwicz scores, GHC requires you to plan to use Bayes' Rule, just as Undominated Dominance does.

So consider case (i) first, where the evidence you will receive will come from partition $\mathcal{E}$. Whatever prior you have, updating on whatever evidence from $\mathcal{E}$ you get will lead to a posterior credence function that has maximal epistemic utility.[34] So, if we apply GHC to prior-plan pairs in this case,

[34] This is because, for any strictly proper measure of epistemic utility $\mathfrak{EU}$, and any world w in $\mathcal{W}$, $\mathfrak{EU}(w, w) \geq \mathfrak{EU}(C, w)$, for all credence functions C, where w is the credence function $w(X) = 1$ if X is true at w and $w(X) = 0$ if X is false at w.

the pairs that maximise the generalised Hurwicz score will have as their prior the same credence function that GHC would have demanded had we applied it only to the choice of prior. So, given the generalised Hurwicz weights above, the prior-plan pair that GHC demands is the one with the uniform distribution, $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, as its prior.

But now consider case (ii), where the evidence you will receive will come not from partition $\mathcal{E}$, but from partition $\mathcal{E}'$. Here, we look at each prior and calculate not only its epistemic utility at each world but the sum of the epistemic utilities of it and the result of updating it by Bayes' Rule on whatever evidence you'll learn at that world. And we measure epistemic utility using the Brier score. Applying GHC to these joint epistemic utilities, we find that the pair that maximises epistemic utility is the one determined by this prior and Bayes' Rule: $(0.333\ldots, 0.429\ldots, 0.238\ldots)$.

So: when we understand ourselves as choosing our prior and ours plans for updating together, as a package, GHC requires that the prior that belongs to the package you pick depends not only on your Hurwicz weights, but also on what evidence you might acquire during your life.[35] Of course, if this is where the argument leads, then we must follow it. But I think it is not. Contrary to the picture that Briggs and I presented, we do not pick our priors and our updating plans together. Rather, we pick our priors and then use those to pick our updating plans. Indeed, when we pick our priors, we pick one of the tools that we will use to pick everything else. We pick a vantage point from which to approach the world and from which to make decisions. We do not pick both that vantage point and our plans for how to pick other things all at once. So, again, I think, it is appropriate to apply GHC only to the choice of priors. We will then see in the following section how to use those priors to pick our posteriors.

6.3 Using priors to pick posteriors

How, then, should we use our prior credences, chosen using GHC, to choose our posterior credences after we acquire some evidence. Here, I turn to recent work by J. Dmitri Gallow (2019). To introduce it, I return to where I started with epistemic utility arguments, in the two papers that Hannes Leitgeb and I wrote in 2010 (Leitgeb & Pettigrew, 2010a,b). In those, we considered two sorts of argument in favour of a genuinely diachronic version of the Bayesian updating norm.

The first ran as follows: Suppose my prior credence function is P, which is a probability function. And suppose I obtain evidence E. How should I respond? Leitgeb and I suggested that evidence E imposes constraints on your posterior credence function, namely, it requires that you assign maximal credence to E—note that this is the same constraint we explored

[35] See (Konek, 2016) for a similar suggestion.

in Section 6.1 above. You should now pick the posterior, among those that satisfy this constraint, that maximises expected epistemic utility from the point of view of your prior.[36] That is, your posterior P' should be

$$P' = \underset{Q \in \mathcal{P}: Q(E)=1}{\arg\min} \left(\sum_{w \in \mathcal{W}} P(w) \mathfrak{EU}(Q, w) \right)$$

where $\mathcal{P}$ is the set of probabilistic credence functions over $\mathcal{F}$.

The first problem with this proposal is that the updating rule it demands is not Bayes' Rule, and indeed it has undesirable features. Leitgeb and I applied the approach using only the Brier score, and Ben Levinstein (2012) offered what I take to be a definitive critique of the rule that results. But it turns out that, for all strictly proper measures of epistemic utility, the approach gives the same updating rule, at least for many priors.[37] And so Levinstein's criticisms apply to the approach as a whole.

The second problem with this approach to updating comes from Dmitri Gallow (2019, 9-10). He notes that we seek a justification for Bayes' Rule that is wholly based on considerations of epistemic utility. That is, we hope not to appeal to any non-telelogical reasoning. But how, Gallow asks, can we justify restricting our possible posteriors to those for which $p(E) = 1$? What justification that appeals only to epistemic utility can we give for that? There is an evidentialist justification, of course. It simply points out that evidence E supports proposition E to the maximal degree, and so someone with evidence E should believe E to the maximum degree, which is credence 1. But what of the epistemic utility theorist? Can they say why we restrict to only those credence functions?

I suspect the following answer gives our best hope. In the dominance argument for Probabilism, we show that, if you do not satisfy the probability axioms, then there is some alternative credence function that has greater epistemic utility at all worlds. We might say that, upon learning E, you have reason not to care about your epistemic utility at worlds where E is false. If that's the case, we might appeal to the following slight tweak of the mathematical theorem on which the dominance argument for Probabilism is based: if credence function C does not satisfy the probability axioms, *or if it does, but* $C(E) < 1$, then there is an alternative credence function $C^\star$ *with* $C^\star(E) = 1$ that has greater epistemic utility than C at all worlds at which E is true. That's why you should assign maximal credence to your evidence.

The problem with this argument is that it takes an inconsistent view of which worlds you should care about. On the one hand, when it applies the

[36] As good card-carrying veritists, and conceiving of ourselves as extending the accuracy-first programme in epistemology that had been initiated by Jim Joyce (1998), Leitgeb and I took accuracy to be the sole source of epistemic utility. But that axiological commitment doesn't make any difference here.

[37] I prove this in Theorem 10 in Section 6.5 below.

tweaked version of the dominance argument to conclude that you should assign maximal credence to your evidence, it says that you should not care about the epistemic utility of a credence function at worlds at which your evidence is false. But, on the other hand, when it comes to calculating the expected epistemic utility of the candidate posterior credence function, you take its epistemic utility at each world, including those at which your evidence is false, weight it by your credence in that world, and include it in the summation that gives your expected epistemic utility. So, ultimately, I think this proposal is not coherent enough to be compelling.

Let's move on, then, to the second argument that Leitgeb and I gave. As before, we argued that learning evidence should lead you to not care about worlds at which it is false; and, as before, we argued that it should lead you to omit those worlds in the decision rules that you follow. But we focused not on the dominance principle but on the principle of maximising expected utility. That is, when we claimed that, when we calculate the expected epistemic utility of a candidate posterior from the point of view of our prior, we should simply omit the summands for worlds at which the evidence is false. Thus, we seek

$$P' = \arg\min_{Q \in \mathcal{P}} \left(\sum_{w \in E} P(w) \mathfrak{EU}(Q, w) \right)$$

And it turns out that, for any strictly proper epistemic utility function, this will be $P'(-) = P(-|E)$, whenever $P(E) > 0$, just as Bayes' Rule requires. Indeed, this is easy to see, since

$$\sum_{w \in E} P(w) \mathfrak{EU}(Q, w) = P(E) \sum_{w \in \mathcal{W}} P(w|E) \mathfrak{EU}(Q, w)$$

Multiplying a function by a constant does not change its maximiser, and since $\mathfrak{EU}$ is strictly proper, $\sum_{w \in \mathcal{W}} P(w|E) \mathfrak{EU}(Q, w)$ is maximised, as a function of Q, at $Q(-) = P(-|E)$.

The problem with this approach, as Gallow points out, is that our usual arguments for choosing by maximising expected utility do carry over to the restricted case in which we sum over only the worlds at which the evidence is true. Fortunately, Gallow also notices an ingenious way to avoid this problem. He suggests that what happens when you learn E is this: you don't exclude the worlds at which E is false from the set over which you run your dominance arguments nor from the set over which you sum to obtain your expectation; rather, you change your epistemic utility function so that:

(i) for a world at which E is true, the value of a credence function is its usual epistemic utility, as measured by a strictly proper scoring rule;

(ii) for a world at which E is false, the value of a credence function is some constant k, the same one for each world.

This has the effect of making us not care about the epistemic utility of a credence function in worlds that our evidence has ruled out. Thus, given an epistemic utility function $\mathfrak{EU}$, a body of evidence E, and a constant k, define the new epistemic utility function $\mathfrak{EU}^k_E$ as follows:

$$\mathfrak{EU}^k_E(P, w) := \begin{cases} \mathfrak{EU}(P, w) & \text{if } w \in E \\ k & \text{if } w \notin E \end{cases}$$

Then it is easy to see that the following functions of Q are maximised at the same point:

(i) $\sum_{w \in E} P(w)\mathfrak{EU}(Q, w)$

(ii) $\sum_{w \in \mathcal{W}} P(w)\mathfrak{EU}^k_E(Q, w)$

So, since (i) is maximised at $Q(-) = P(-|E)$, so is (ii). Borrowing it from Gallow, then, this is our argument for updating using Bayes' Rule.

6.4 Tying up loose ends

There are two loose ends to tie up before this gives us a satisfactory account of credal updating. First, you might worry that, while your prior was an appropriate probability function to use to calculate expectations before the new evidence arrived, it is no longer appropriate now that the new evidence is here, and so it isn't rationally required to update to whatever maximises expected epistemic value from its point of view. Second, you might worry that this solution requires us to choose using GHC when we have no evidence and no prior, and then using expected utility when we have a prior and some evidence—and you might wonder what could motivate this difference. In this section, I answer these worries.

6.4.1 Maximising expected epistemic utility from whose point of view?

The first objection usually runs as follows. At an earlier time, you have a credence function C and a total body of evidence E; and we suppose C is a rational response to E. But then your evidence changes, because you learn something new—specifically, you learn E'. And now you want to know how to respond to this. According to Gallow's approach, which I am adopting here, you call on your credence function C and you calculate the expected epistemic utilities for the various candidate posteriors relative to that and using an epistemic utility function $\mathfrak{EU}^k_{E'}$, where k is some constant and $\mathfrak{EU}$ is strictly proper. But surely your credence function C has no authority any longer? After all, while it respects your evidence E at the earlier time, it does not respect your new evidence E'; it gives positive credence to worlds at which E' is false. So why should we pay any attention to the expected epistemic utilities calculated from its point of view?

I think the first thing to say is that the epistemic utility theorist has no notion of respecting the evidence that does not flow ultimately from considerations of epistemic utility. So it isn't obvious, before you look at the epistemic utilities, whether or not your current credence function does or does not respect the evidence, even supposing that notion has any meaning. But Gallow's approach does allow us to see that there is something amiss with your prior as a response to this new evidence. And indeed it gives a recommendation for how to correct that error. It is irrational, we might suppose, to have a credence function that expects another to be better. So, we might suppose that C expects itself to be best relative to the epistemic utility function $\mathfrak{EU}^k_E$. But now you have evidence E' and your new epistemic utility function is $\mathfrak{EU}^k_{E'}$. And now C no longer passes the test (unless $C(E') = 0$ or $C(E') = 1$). But it does tell you how to respond—you should adopt whichever credence function does now maximise epistemic utility function from the point of view of C and relative to the epistemic utility function $\mathfrak{EU}^k_{E'}$. And, as Gallow shows, that is $C(-|E')$.

6.4.2 In favour of maximising expected utility

The second objection is due to Sophie Horowitz (2017, Footnote 24). As I read it, she presents two challenges to the sort of approach I champion here. The first, to say why we should use different decision rules at each stage; the second, to say why these principles, in particular—i.e. GHC to pick your priors, and then the expected utility rule thereafter.

I think the first is reasonably easy to answer. In Section 4, we noted that we can classify decision rules by the ingredients they take as inputs. For instance, some take as inputs only a representation of the possibilities and a representation of the agent's conative attitudes. Wald's Maximin principle is like this. We represent the possibilities by possible worlds and the conative attitudes by utilities. And that's all that Maximin requires to specify the rationally permissible choices. Others take a representation of the possibilities, a representation of the conative attitudes, and a representation of the risk attitudes. The original Hurwicz criterion is like this, as is our generalised version. Possibilities are again represented possible worlds, conative attitudes by utilities, and risk attitudes by Hurwicz weights. Others still take possibilities, conative attitudes, and doxastic attitudes, the latter represented by credence functions. Maximising expected utility is an instance of this. And some add attitudes to risk to the conative and doxastic attitudes. Lara Buchak's risk-weighted expected utility theory is such a decision rule.

Now, at the beginning of your epistemic life, before you have assigned credences, you cannot appeal to a principle that requires as an input your doxastic state—after all, you don't have one. And we argued above that, in such a situation, you should adopt GHC. However, once you have acquired a doxastic state, new decision rules become available. In particular, those that require a doxastic state. For instance, maximising expected utility or maximising risk-weighted expected utility, among others. Which should you choose? I'll build on work by Martin Peterson and Kenny Easwaran to argue that you should choose by maximising expected utility relative to your credences.

Recall how we argued for GHC above. We began with just the possible worlds and the utility function over those worlds, and laid down axioms governing your preference orderings $\preceq$. And we showed that, for any ordering that satisfies those axioms, there are Hurwicz weights $\Lambda = (\lambda_1, \ldots, \lambda_n)$ such that, for any acts a and a',

$$a \preceq a' \text{ iff } H^\Lambda(a) \leq H^\Lambda(a')$$

This time, we are not in the same situation. As well as the possible worlds and the utility function, we also have a probabilistic credence function P defined over the possible worlds. Now we wish to show that, in this case, your preferences should be as follows:

$$a \preceq a' \text{ iff } \sum_{w \in \mathcal{W}} P(w)a(w) \leq \sum_{w \in \mathcal{W}} P(w)a'(w)$$

We call this biconditional *the expected utility principle*. We want to give what is sometimes called an *ex ante* justification for this principle.

This is not the usual sort of justification that is given for that norm. In Savage's original work, he proceeded as follows (Savage, 1954). He assumed that you have a preference ordering $\preceq$ over a rich set of acts; he imposed rational requirements on that preference ordering; and he showed that, if your preference ordering satisfies these constraints, there is a probability function that assigns a credence $P(w)$ to each world w in $\mathcal{W}$, and a utility function that, for each option a, assigns a utility $a(w)$ to each world w, such that

$$a \preceq a' \text{ iff } \sum_{w \in \mathcal{W}} P(w)a(w) \leq \sum_{w \in \mathcal{W}} P(w)a'(w)$$

This is Savage's representation theorem. It provides what is called an *ex post* justification of the expected utility principle. It begins with your preference ordering and extracts from it credences and utilities for which expected utility orders the acts just as your preference ordering does. But a natural question arises: are the credences and utilities extracted from your preference ordering in this way the credences and utilities that you assign? Are they your credences and utilities? For some philosophers and for many economists, the answer is, by definition, yes. They take a representation theorem like Savage's to say what it means to have a particular credence function and a particular utility function: it is simply to have a preference ordering that satisfies Savage's axioms and that orders acts by their expected utilities relative to that credence function and utility function. Others will say that credences, utilities, and preferences have separate, independent existences; it is possible to have credences and utilities that do not match your preferences in the way the expected utility principle demands. But they might also say that, when they don't, you are irrational. That is, the expected utility principle is not part of a definition of what it means to have certain credences and utilities; rather, it is a norm that governs how your credences, utilities, and preference ordering should relate. This is the view I take here. But Savage's representation theorem provides no route to it. Instead of a list of conditions on a preference ordering alone and a theorem that shows that, for any preference ordering that satisfies those conditions, there exist a credence function and utility function that, together with the preference ordering, satisfy the expected utility principle, we need a list of conditions on the relationship between a preference ordering, a credence function, and a utility function and a theorem that shows that any three that satisfy those conditions together satisfy the expected utility principle. Fortunately, such justifications already exist. I am aware of two: one due to Martin Peterson (2004); the other, more general one, due to Kenny Easwaran (2014).

Some of the conditions we will impose on the relationship between preferences, credences, and utilities are familiar; others less so. Recall Coarse Grain Indifference (A8) from Section 4.3 above. To state that, we had to talk of coarse-grainings of our individual's set of possible worlds. To state our conditions here, we need to talk about fine-grainings of our individual's set of possible worlds, and then fine-grainings of their probability and utility functions as well.

- A coarse-graining of $\mathcal{W}$ is a set $\mathcal{W}^*$ together with a surjective function $h : \mathcal{W} \to \mathcal{W}^*$.
- A fine-graining of $\mathcal{W}$ is a set $\mathcal{W}'$ together with a surjective function $h : \mathcal{W}' \to \mathcal{W}$.

Suppose $\mathcal{W} = \{w_1, w_2, w_3\}$, where Biden wins at w_1, Trump at w_2, and Jorgensen at w_3.

- Here is a coarse-graining of $\mathcal{W}$: $\mathcal{W}^* = \{w_1^*, w_2^*\}$, where Biden wins at w_1^* and Biden loses at w_2^*. So $h(w_1) = w_1^*$ and $h(w_2) = h(w_3) = w_2^*$.
- Here is a fine-graining of $\mathcal{W}$: $\mathcal{W}' = \{w'_1, w'_2, w'_3, w'_4\}$, where Biden wins and is left-handed at w'_1, Biden wins and is right-handed at w'_2, Trump wins at w'_3, and Jorgensen wins at w'_4. So $h(w'_1) = h(w'_2) = w_1$, $h(w'_3) = w_2$, and $h(w'_4) = w_3$.

See Figure 7 for an illustration.

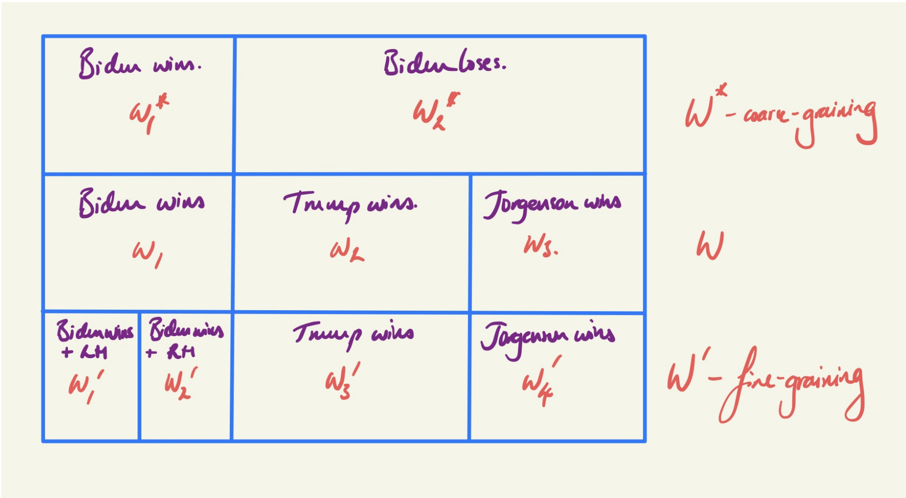

Figure 7: Our example of fine- and coarse-grainings.

Given a probability function P on $\mathcal{W}$, and a fine-graining $\mathcal{W}', h$ of $\mathcal{W}$, we say that a probability function P' on $\mathcal{W}'$ is a fine-graining of P if, for all w in $\mathcal{W}$,

$$\sum_{\substack{w' \in \mathcal{W}' \\ h(w') = w}} P'(w') = P(w)$$

That is, the probability of a world in $\mathcal{W}$ is the sum of the probabilities assigned to its fine-grainings in $\mathcal{W}'$.

Given an option a that is defined on $\mathcal{W}$, and a fine-graining $\mathcal{W}', h$ of $\mathcal{W}$, we say that an option a' on $\mathcal{W}'$ is the fine-graining of a if, for all w' in $\mathcal{W}'$, $a'(w') = a(h(w'))$. That is, the utility of a' at a fine-graining of world w is just the utility of a at w.

Here are the conditions. They are imposed not only on the individual's preference ordering $\preceq_{\mathcal{W},P}$ on options defined on $\mathcal{W}$ and where their credence function is P, but also on their preference orderings $\preceq_{\mathcal{W}',P'}$ on fine-grainings of options of those options and where their credence function P' on $\mathcal{W}'$ is a fine-graining of P.

(A1) **Reflexivity** $\preceq_{\mathcal{W}',P'}$ is reflexive.

(A2) **Transitivity** $\preceq_{\mathcal{W}',P'}$ is transitive.

As before, these are uncontroversial in this context, so I'll say no more about them.

(A3*) **Limit Continuity** If $a_1, a_2, \ldots$ tends to a and $a' \prec_{\mathcal{W}',P'} a_i$ for all i, then $a' \preceq_{\mathcal{W}',P'} a$.

This is a slightly different continuity axiom from the one we imposed above, but it is no less natural.

(A4) **Weak Dominance**

(i) If $a(w) \leq a'(w)$ for all w in $\mathcal{W}$, then $a \preceq_{\mathcal{W}',P'} a'$

(ii) If $a(w) < a'(w)$ for all w in $\mathcal{W}$, then $a \prec_{\mathcal{W}',P'} a'$.

Again, this is uncontroversial.

(A8*) **Probabilistic Fine Grain Indifference** Suppose $\mathcal{W}'$ and $h : \mathcal{W}' \to \mathcal{W}$ is a fine-graining of $\mathcal{W}$, suppose P' defined on $\mathcal{W}'$ is a fine-graining of P defined on $\mathcal{W}$, and suppose options a' and b' defined on $\mathcal{W}'$ are fine-grainings of options a and b defined on $\mathcal{W}$, respectively. Then $a \sim_{\mathcal{W},P} b$ iff $a' \sim_{\mathcal{W}',P'} b'$.

Why require this? After all, we rejected the coarse-graining version when we offered our characterization of GHC above. But, in fact, it's for that very reason that we should accept it here. We rejected it in our characterization of the decision rule you should use when you have no probabilities and no evidence because it conflicts with Permutation Indifference and Strong Dominance, and because Permutation Indifference is a more compelling principle in that situation. But now we are considering our agent in a different situation. They might still have no evidence, but they do have probabilities—either prior probabilities that they set by appealing to GHC, or probabilities obtained from those priors using Bayes' Rule. And so Permutation Indifference is no longer appropriate. Once you have set your credences in the possible worlds, it is no longer reasonable for us to demand that you are indifferent between options that differ only in which worlds receive which utilities, but do not differ in the utilities assigned. So, having dropped Permutation Indifference, we clear away the obstacle that prevented us from adopting the other very compelling sort of principle that we wished to adopt above, but which were prevented from adopting by the threat of inconsistency. And indeed, we can now see why Strong Dominance, Permutation Indifference, and Coarse Grain Indifference are all so plausible despite being inconsistent. Strong Dominance is a requirement at all times; Permutation Indifference is a requirement only when you have no probabilities and no evidence; and the various versions of Grain Indifference are required only when you do have probabilities, whether or not you also have evidence.

(A10) **Trade-Off Indifference** If, for two possible worlds w'_i, w'_j in $\mathcal{W}'$,

(i) $P(w'_i) = P(w'_j)$,

(ii) $a(w'_i) - b(w'_i) = b(w'_j) - a(w'_j)$, and

(iii) $a(w'_k) = b(w'_k)$ for all $w'_k \neq w'_i, w'_j$,

then $a \sim_{\mathcal{W}',P'} b$.

This says that, if the utility of one option a exceeds the utility of another b at one world w'_i, but the utility of b exceeds the utility of a by the same amount at another world w'_j, and if you think each of those worlds equally likely, and if a and b have the same utility at all other worlds, then you should be indifferent between a and b. That is, b compensates perfectly for its lower utility at w'_i by having the higher utility at w'_j. So, for instance, if $\mathcal{W}' = \{w'_1, w'_2\}$, and you assign equal credence to each world, then Trade-Off Indifference demands $(0,8) \sim (1,7) \sim (2,6) \sim (3,5) \sim (4,4)$.

Note, however, that this demands a sort of risk-neutrality. After all, we typically expect that the risk-averse and the risk-inclined will agree on their evaluation of a risk-free option like $(4,4)$. But they will disagree on an option like $(0,8)$ or $(1,7)$ that involves risk. But Trade-Off Indifference rules that out: both must evaluate $(0,8)$ in the same way, namely, as equivalent to the risk-free option $(4,4)$, which they both treat in the same way.

Surely, then, this goes against the Jamesian permissivism about attitudes to epistemic risk that has been the central driving force of my account so far? I think not. If you were to use a risk-sensitive decision rule to pick your priors, and then a risk-sensitive decision rule also to pick your posteriors using your priors, you would double count your attitudes to risk. This is the answer to Sophie Horowitz's challenge to this brand of Jamesian epistemology. Horowitz asks why we should use a risk-sensitive decision rule at the beginning of our epistemic life, and then a risk-neutral one thereafter. The answer is that we encode our attitudes to risk entirely in the decision rule we use to pick our priors. Having picked them using that rule, we need not and indeed should not include our attitudes to risk also in the decision rule we use when we appeal to those priors to pick our posteriors. That would be double counting. It would be as if in the practical case, we were to encode our attitudes to risk entirely in our decision rule, and then also encode them a second time in our utility function.

Then we have the following theorem:[38]

Theorem 8 *Suppose $\prec, \preceq$ satisfy (A1), (A2), (A3*), (A4), (A8*), and (A10). Then*

$$a \preceq a' \text{ iff } \sum_{w \in \mathcal{W}} P(w)a(w) \leq \sum_{w \in \mathcal{W}} P(w)a'(w)$$

And that completes our argument. When you come to a decision armed with a doxastic state as well as a conative state, you should choose by maximising your expected utility. And, as we have seen above, thanks to Gallow's argument, that requires us to update using Bayes' Rule.

[38]The proof proceeds in the same way as the proofs in (Peterson, 2004) and (Easwaran, 2014).

6.5 Appendix: proofs

Theorem 9 *Suppose $\Lambda = (\lambda_1, \lambda_2, \lambda_3)$ is a sequence of Hurwicz weights, $\mathfrak{s}$ is a strictly proper scoring rule, $\mathcal{W} = \{w_1, w_2, w_3\}$, and $E = \{w_1, w_2\}$. Then:*

(i) *If $\lambda_2 \geq \lambda_1$, then the credence function $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$ uniquely maximises $H^\Lambda(\mathfrak{EU}(-))$ among credence functions Q with $Q(E) = 1$.*

(ii) *If $\lambda_2 < \lambda_1$, then $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$ does not maximise $H^\Lambda(\mathfrak{EU}(-))$ among credence functions Q with $Q(E) = 1$.*

Proof. Suppose $\mathfrak{s}$ is a strictly proper scoring rule, φ is the associated entropy function, and $\mathfrak{D}_\mathfrak{s}$ is the corresponding divergence, as defined in the proof of Theorem 7 above. Then $(x, 1-x, 0)$ and $(1-x, x, 0)$ maximise $H^\Lambda(\mathfrak{EU}(-))$ among C with $C(E) = 1$ iff $x \geq 1-x$ and $(x, 1-x, 0)$ minimises $\mathfrak{D}_\mathfrak{s}((\lambda_1, \lambda_2, \lambda_3), (x, 1-x, 0))$ among those $x > 1-x$. Then, as in the proof of Theorem 7, we appeal to the KKT conditions. We let

$$\begin{aligned}
f(x) &= \mathfrak{D}_\mathfrak{s}((\lambda_1, \lambda_2, \lambda_3), (x, 1-x, 0)) \\
&= [\varphi(\lambda_1) - \varphi(x) - \varphi'(x)(\lambda_1 - x)] + \\
&\quad [\varphi(\lambda_2) - \varphi(1-x) - \varphi'(1-x)(\lambda_2 - (1-x))] + \\
&\quad [\varphi(\lambda_3) - \varphi(0) - \varphi'(0)(\lambda_3 - 0)] \\
&\quad [\varphi(\lambda_1 + \lambda_2) - \varphi(1) - \varphi'(1)((\lambda_1 + \lambda_2) - 1)] + \\
&\quad [\varphi(\lambda_1 + \lambda_3) - \varphi(x) - \varphi'(x)((\lambda_1 + \lambda_3) - x)] + \\
&\quad [\varphi(\lambda_2 + \lambda_3) - \varphi(1-x) - \varphi'(1-x)((\lambda_2 + \lambda_3) - (1-x))]
\end{aligned}$$

$$g(x) = -x + \frac{1}{2}$$

Then, by the KKT theorem, $(x, 1-x, 0)$ is a minimiser iff there is $\mu \geq 0$ such that

$$\frac{d}{dx}\mathfrak{D}_\mathfrak{s}((\lambda_1, \lambda_2, \lambda_3), (x, 1-x, 0)) - \mu = 0$$

Now

$$\begin{aligned}
\frac{d}{dx}[\varphi(\lambda_1) - \varphi(x) - \varphi'(x)(\lambda_1 - x)] &= \varphi''(x)(x - \lambda_1) \\
\frac{d}{dx}[\varphi(\lambda_2) - \varphi(1-x) - \varphi'(x)(\lambda_2 - (1-x))] &= \varphi''(1-x)(\lambda_2 - (1-x)) \\
\frac{d}{dx}[\varphi(\lambda_3) - \varphi(0) - \varphi'(0)(\lambda_3 - 0)] &= 0 \\
\frac{d}{dx}[\varphi(\lambda_1 + \lambda_2) - \varphi(1) - \\
\varphi'(1)((\lambda_1 + \lambda_2) - 1)] &= 0 \\
\frac{d}{dx}[\varphi(\lambda_1 + \lambda_3) - \varphi(x) - \\
\varphi'(x)((\lambda_1 + \lambda_3) - x))] &= \varphi''(x)(x - (\lambda_1 + \lambda_3)) \\
\frac{d}{dx}[\varphi(\lambda_2 + \lambda_3) - \varphi(1-x) - \\
\varphi'(1-x)((\lambda_2 + \lambda_3) - (1-x))] &= \varphi''(1-x)((\lambda_2 + \lambda_3) - (1-x))
\end{aligned}$$

So

$$\begin{aligned}
0 &= \frac{d}{dx}\mathfrak{D}_{\mathfrak{s}}((\lambda_1, \lambda_2, \lambda_3), (1/2, 1/2, 0)) - \mu \\
&= \varphi''(1/2)(1/2 - \lambda_1) + \varphi''(1/2)(\lambda_2 - 1/2) + \\
&\quad \varphi''(1/2)(1/2 - (\lambda_1 + \lambda_3)) + \varphi''(1/2)((\lambda_2 + \lambda_3) - 1/2) - \mu \\
&= 2\varphi''(1/2)(\lambda_2 - \lambda_1) - \mu \\
&\text{iff} \\
\mu &= 2\varphi''(1/2)(\lambda_2 - \lambda_1)
\end{aligned}$$

Now, since φ is strictly convex, φ'' is strictly positive, and in particular $\varphi''(1/2) > 0$. So, if $\lambda_2 \geq \lambda_1$, then let $\mu \geq 0$, giving (i). And, if $\lambda_2 < \lambda_1$, then $\mu < 0$, giving (ii). □

Theorem 10 *Suppose $\mathcal{W} = \{w_1, w_2, w_3\}$ and $E = \{w_1, w_2\}$. Suppose $\mathfrak{EU}$ is strictly proper and generated by $\mathfrak{s}$. And suppose $P = (p_1, p_2, p_3)$ is your prior credence function. Then define P' as follows:*

$$\begin{aligned}
p'_1 &= p_1 + \frac{1}{2}p_3 \\
p'_2 &= p_2 + \frac{1}{2}p_3 \\
p'_3 &= 0
\end{aligned}$$

Then P' is the credence function that maximises $\text{Exp}^P(\mathfrak{EU}(-))$ among credence functions Q with $Q(E) = 1$.

Proof. The posterior that maximises $\text{Exp}^P(\mathfrak{EU}(-))$ among credence functions that assign credence 1 to E has the form $(x, 1-x, 0)$. Let $\mathfrak{s}$ be a strictly proper scoring rule. Then, writing $\mathfrak{s}_i(-)$ for $\mathfrak{s}(i,-)$ for the sake of brevity,

$$\begin{aligned}
&\text{Exp}^P(\mathfrak{EU}((x, 1-x, 0))) = \\
&p_1[\mathfrak{s}_1(x) + \mathfrak{s}_0(1-x) + \mathfrak{s}_0(0) + \mathfrak{s}_1(x + (1-x)) + \mathfrak{s}_1(x+0) + \mathfrak{s}_0((1-x)+0)] + \\
&p_2[\mathfrak{s}_0(x) + \mathfrak{s}_1(1-x) + \mathfrak{s}_0(0) + \mathfrak{s}_1(x + (1-x)) + \mathfrak{s}_0(x+0) + \mathfrak{s}_1((1-x)+0)] + \\
&p_3[\mathfrak{s}_0(x) + \mathfrak{s}_0(1-x) + \mathfrak{s}_1(0) + \mathfrak{s}_0(x + (1-x)) + \mathfrak{s}_1(x+0) + \mathfrak{s}_1((1-x)+0))]
\end{aligned}$$

Now, ignore the constant terms, since they do not affect the minima; and group terms together. Then we wish to maximise:

$$\mathfrak{s}_1(x)(2p_1+p_3) + \mathfrak{s}_0(1-x)(2p_1+p_3) + \mathfrak{s}_1(1-x)(2p_2+p_3) + \mathfrak{s}_0(x)(2p_2+p_3)$$

Now, divide through by 2, which again doesn't affect the minimisation. Then we wish to maximise:

$$\mathfrak{s}_1(x)(p_1+\frac{1}{2}p_3) + \mathfrak{s}_0(1-x)(p_1+\frac{1}{2}p_3) + \mathfrak{s}_1(1-x)(p_2+\frac{1}{2}p_3) + \mathfrak{s}_0(x)(p_2+\frac{1}{2}p_3)$$

But $(p_1 + \frac{1}{2}p_3) + (p_2 + \frac{1}{2}p_3) = p_1 + p_2 + p_3 = 1$. So, since $\mathfrak{s}$ is strictly proper,

$$\mathfrak{s}_1(x)(p_1+\frac{1}{2}p_3) + \mathfrak{s}_0(x)(p_2+\frac{1}{2}p_3)$$

is maximised, as a function of x at $x = p_1 + \frac{1}{2}p_3 = p'_1$. And

$$\mathfrak{s}_1(1-x)(p_2+\frac{1}{2}p_3) + \mathfrak{s}_0(1-x)(p_1+\frac{1}{2}p_3)$$

is maximised, as a function of x, at $1 - x = p_2 + \frac{1}{2}p_3 = p'_2$. So

$$\mathfrak{s}_1(x)(p_1+\frac{1}{2}p_3) + \mathfrak{s}_0(1-x)(p_1+\frac{1}{2}p_3) + \mathfrak{s}_1(1-x)(p_2+\frac{1}{2}p_3) + \mathfrak{s}_0(x)(p_2+\frac{1}{2}p_3)$$

is maximised, as a function of x at $x = p_1 + \frac{1}{2}p_3 = p'_1$. So $\text{Exp}^P(\mathfrak{EU}((x, 1-x, 0)))$ is maximised at $P' = (p'_1, p'_2, p'_3)$, as required. □

In brief...

In the preceding two chapters, I asked how we should use epistemic utilities to pick our priors. In this chapter, I turned my attention to picking posteriors. In Sections 6.1-6.2, I considered and rejected a range of possible approaches. In Sections 6.3-6.4, I settled on an argument by Dmitri Gallow (2019) that we should pick our posteriors by maximising expected epistemic utility from the point of view of our priors, which entails that we should obtain our posteriors from our priors by conditioning on our evidence, just as Bayes' Rule demands.

At first, this approach looks unpromising. After all, we have assumed that all epistemic utility functions are strictly proper, and that means that each probabilistic credence function expects itself to be best. But in that case, your probabilistic prior will demand that you adopt it as your posterior, since doing so maximises expected epistemic utility from its point of view; and doing that ignores the evidence you've acquired. However, Gallow argues persuasively that our epistemic utility function should be strictly proper only at the beginning of our epistemic life. After that, as we acquire evidence, it should change. At a given time, our epistemic utility function should treat worlds at which our evidence is true differently from worlds at which it is false. At worlds at which it is true, the epistemic utility of a credence function should be whatever our strictly proper prior epistemic utility function says it is; but at worlds at which our evidence is false, the epistemic utility of a credence function should just be some constant—the same for each credence function. This is because, once we acquire evidence, we no longer care about the epistemic utility of our crednece function at worlds that our evidence has ruled out. This adapted epistemic utility function encodes that attitude to those worlds, since it treats all credence functions in the same way at those worlds. Buildling on an argument that Hannes Leitgeb and I gave, Gallow then notes that, if epistemic utility is measured in this way, the posterior that maximises expected epistemic utility from the point of view of our prior is precisely the one obtained by conditioning our prior on our evidence (Leitgeb & Pettigrew, 2010b).

Now, you might wonder why we should choose posteriors using a risk-neutral decision rule like the expected utility rule. My answer is that, if we use a risk-sensitive decision rule like GHC to pick priors, and encode our attitudes to epistemic risk in the Hurwicz weights we use for that choice, and then we use risk-sensitive rule to pick posteriors, and encode our attitudes to epistemic risk in that rule, we would double count our attitudes to epistemic risk. It would be as bad as if we were to encode our attitudes to risk first in our utilities in the pragmatic case, and then again in our decision rule.

7 Answering objections

In the preceding six chapters, I have offered my positive argument for permissivism about epistemic rationality; more specifically, permissivism about our credal response to evidence. In this chapter, I defend my argument against objections that have been raised to permissivism more generally.

7.1 Permissive rationality and deference

One type of argument against permissivism about epistemic rationality turns on the claim that rationality is worthy of deference. The argument begins with a precise version of this claim, stated as a norm that governs credences. It proceeds by showing that, if epistemic rationality is permissive, then it is sometimes impossible to meet the demands of this norm. Taking this to be a reductio, the argument concludes that rationality cannot be permissive. I know of two versions of the argument, one due to Daniel Greco and Brian Hedden, and one due to Ben Levinstein (Greco & Hedden, 2016; Levinstein, 2017). I'll begin with Levinstein's, since his version of the deference norm fixes some problems with Greco and Hedden's. I'll consider David Thorstad's response to Greco and Hedden's argument, which would also work against Levinstein's argument were it to work at all (Thorstad, 2019). But I'll conclude that, while it provides a crucial insight, it doesn't quite work, and I'll offer my own alternative response. And finally I'll return to Greco and Hedden's argument for their norm and explain why I think it fails.

Roughly speaking, you defer to someone on an issue if, upon learning their attitude to that issue, you adopt it as your own. So, for instance, if you ask what I'd like to eat for dinner tonight, and I say that I defer to you on that issue, I'm saying that I will want to eat whatever I learn you want to eat. That's a case of deferring to someone else's preferences—it's a case where we defer *conatively* to them. Here, we are interested in cases in which we defer to someone else's beliefs—that is, where we defer *doxastically* to them. Thus, I defer doxastically to my radiographer on the issue of whether I've got a broken finger if I commit to adopting whatever credence they announce as their diagnosis. By analogy, we sometimes say that we defer doxastically to a feature of the world if we commit to setting our credence in some way that is determined by that feature of the world. So, I might defer doxastically to a particular computer simulation model of Earth's oceans on the issue of sea level rise by 2030 if I commit to setting my credence in a rise of 10cm to whatever probability that model reports when I run it repeatedly while perturbing its parameters and initial conditions slightly around my best estimate of their true values.

In philosophy, there are a handful of well-known theses that turn on the claim that we are required to defer doxastically to this individual or that feature of the world—and we're required to do it on all matters. For instance, Bas van Fraassen's Reflection Principle says that you should defer doxastically to your future self on all matters (van Fraassen, 1984, 1995; Briggs, 2009). That is, for any proposition X, conditional on your future self having credence r in X, you should have credence r in X. In symbols:

$$C(X \mid \text{my credence in } X \text{ at future time } t \text{ is } r) = r$$

And David Lewis' Principal Principle says that you should defer to the objective chances on all doxastic matters by setting your credences to match the probabilities that they report (Lewis, 1980, 1994; Hall, 1994; Thau, 1994). That is, for any proposition X, conditional on the objective chance of X being r, you should have credence r in X. In symbols:

$$C(X \mid \text{the objective chance of } X \text{ now is } r) = r$$

Notice that, in both cases, there is a single expert value to which you defer on the matter in question. At time t, you have exactly one credence in X, and the Reflection Principle says that, upon learning that single value, you should set your credence in X to it. And there is exactly one objective chance of X now, and the Principal Principle says that, upon learning it, you should set your credence in X equal to it. You might be uncertain about what that single value is, but it is fixed and unique. So this account of deference does not cover cases in which there is more than one expert. For instance, it doesn't obviously apply if I defer not to a specific climate model, but to a group of them. In those cases, there is usually no fixed, unique value that is the credence they all assign to a proposition. So principles of the same form as the Reflection Principle or the Principal Principle do not say what to do if you learn one of those values, or some of them, or all of them. This problem lies at the heart of the deference argument against permissivism. Those who make the argument think that deference to groups should work in one way; those who defend permissivism against it think it should work in some different way.

As I mentioned above, the deference argument begins with a specific, precise norm that is said to govern the deference we should show to rationality. The argument continues by claiming that, if rationality is permissive, then it is not possible to satisfy this norm. Here is the norm as Levinstein (2017, 361) states it, where $C \in \mathbf{R}_E$ means that C is in the set $\mathbf{R}_E$ of rational responses to evidence E.

Deference to Rationality (DR) Suppose:

(i) C is your credence function;

(ii) E is your total evidence;

(iii) $C(C \notin \mathbf{R}_E) = 1$;

(iv) C' is a probabilistic credence function;

(v) $C(C' \in \mathbf{R}_E) > 0$;

then rationality requires

$$C(-|C' \in \mathbf{R}_E) = C'(-|C' \in \mathbf{R}_E)$$

That is, if you are certain that your credence function is not a rational response to your total evidence, then, conditional on some alternative probabilistic credence function being a rational response to that evidence, you should set your credences in line with that alternative once you've conditioned that on the assumption that it is a rational response to your original total evidence.

Notice, first, that Levinstein's principle is quite weak. It does not say of just anyone that they should defer to rationality. It says only that, if you are in the dire situation of being certain that you are yourself irrational, then you should defer to rationality. If you are sure you're irrational, then your conditional credences should be such that, were you to learn of a credence function that it's a rational response to your evidence, you should fall in line with the credences that it assigns conditional on that same assumption that it is rational. Restricting its scope in this way makes it more palatable to permissivists who will typically not think that someone who is already pretty sure that they are rational must switch credences when they learn that there are alternative rational responses out there.

Notice also that you need only show such deference to rational credence functions that satisfy the probability axioms. This restriction is essential, for otherwise DR will force you to violate the probability axioms yourself. After all, if $C(-)$ is probabilistic, then so is $C(-|X)$ for any X with $c(X) > 0$. Thus, if $C'(-|C' \in \mathbf{R}_E)$ is not probabilistic, and C defers to C' in the way Levinstein's principle demands, then $C(-|C' \in \mathbf{R}_E)$ is not probabilistic, and thus neither is C.

Now, suppose:

- C is your credence function;
- E is your total evidence;
- C' and C'' are probabilistic credence functions with

$$C'(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E) \neq C''(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E)$$

  That is, C' and C'' are distinct and remain distinct even once they become aware that both are rational responses to E;

- $C(C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E) > 0$. That is, you give some credence to both of them being rational responses to E;
- $C(C \notin \mathbf{R}_E) = 1$. That is, you are certain that your own credence function is not a rational response to E.

Then, by DR,

- $C(-|C' \in \mathbf{R}_E) = C'(-|C' \in \mathbf{R}_E)$
- $C(-|C'' \in \mathbf{R}_E) = C''(-|C'' \in \mathbf{R}_E)$

Thus, conditioning both sides of the first identity on $C'' \in \mathbf{R}_E$ and both sides of the second identity on $C' \in \mathbf{R}_E$, we obtain

- $C(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E) = C'(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E)$
- $C(-|C'' \in \mathbf{R}_E \;\&\; C' \in \mathbf{R}_E) = C''(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E)$

But, by assumption, $C'(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E) \neq C''(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E)$. So DR cannot be satisfied.

One thing to note about this argument: if it works, it establishes not only that there can be no two different rational responses to the same evidence, but also that it is irrational to be anything less than certain of this. After all, what is required to derive the contradiction from DR is not that there are two probabilistic credence functions C' and C'' such that $C'(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E) \neq C''(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E)$ *that are both rational responses to E*. Rather, what is required is only that there are two probabilistic credence functions C' and C'' with $C'(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E) \neq C''(-|C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E)$ *that you think might both be rational responses to E*—that is, $C(C' \in \mathbf{R}_E \;\&\; C'' \in \mathbf{R}_E) > 0$. The reason is that Levinstein's principle applies to your prior. It doesn't tell you directly how you should update upon learning $C' \in \mathbf{R}_E$. Rather, it tells you what conditional credences you should have before learning that. And so, in order to derive the contradiction from DR, Levinstein does not need to appeal to what happens if you actually learn $C' \in \mathbf{R}_E$. He needs only to look at the conditional credences that you have before learning that. And so at no point does he need to assume that there actually are two credence functions C' and C'' that are both rational and that disagree even after both condition on them both being rational. He just needs to assume that you assign a positive probability to this possibility. And so it is assigning that positive probability that his argument shows to be irrational. Now, the conclusion that it is irrational to even entertain permissivism strikes me as too strong, but perhaps Levinstein would be happy to accept it.

Let's turn, then, to a more substantial worry, given compelling voice by David Thorstad (2019): DR is too strong because the deontic modality that features in it is too strong. As I hinted above, the point is that the form of the deference principles that Levinstein and Greco and Hedden use is borrowed from cases—such as the Reflection Principle and the Principal Principle—in which there is just one expert value, though it might be unknown to you. In those cases, it is appropriate to say that, upon learning the single value and nothing more, you are *required* to set your credence in line with it. But, unless we simply beg the question against permissivism and assume there is a single rational response to every body of evidence, this isn't our situation. Rather, it's more like the case where you defer to a group of experts, such as a group of climate models. And in this case, Thorstad says, it is inappropriate to *demand* that you set your credence in line with an expert's credence when you learn what it is. Rather, it is at most appropriate to *permit* you to do that. That is, Levinstein's principle should not say that rationality *requires* your credence function to assign the conditional credences stated in its consequent; it should say instead that rationality *permits* it.

Thorstad motivates his claim by drawing an analogy with a moral case that he describes. Suppose you see two people drowning. They're called John and James, and you know that you will be able to save at most one. So the actions available to you are: (i) *save John*, (ii) *save James*, and (iii) *save neither*. And the moral actions are: (i) *save John*, and (ii) *save James*. But now consider a deference principle governing this situation that is analogous to DR: it demands that, upon learning that it is moral to save James, you must do that; and upon learning that it is moral to save John, you must do that. From this, we can derive a contradiction in a manner somewhat analogous to that in which we derived the contradiction from DR above: if you learn both that it is moral to save John and moral to save James, you should do both; but that isn't an available action; so moral permissivism must be false. But I take it no moral theory will tolerate that in this case—no moral theory will be able to break the symmetry between saving James and saving John, and so uniqueness will require us to save neither, which no reasonable moral theory will allow.[39] So, Thorstad argues, there must be something wrong with the moral deference principle; and, by analogy, there must be something wrong with the analogous doxastic principle, DR.

Thorstad's diagnosis is this: the correct deference principle in the moral case should say: upon learning that it is moral to save James, you may do that; upon learning that it is moral to save John, you may do that. You thereby avoid the contradiction, and moral permissivism is safe. Similarly, the correct doxastic deference principle is this: upon learning that a credence function is rational, it is permissible to defer to it. In Levinstein's framework, the following is rationally permissible, not rationally

[39]Perhaps there is a moral theory on which saving one but not the other would involve treating the two so unequally that it would be immoral. I would not count such a theory as reasonable.

mandated:

$$C(-|C' \in \mathbf{R}_E) = C'(-|C' \in \mathbf{R}_E)$$

I think Thorstad's example is extremely illuminating, but for reasons rather different from his. Recall that a crucial feature of Levinstein's version of the deference argument against permissivism is that it applies only to people who are certain that their current credences are irrational. If we add the analogous assumption to Thorstad's case, his verdict is less compelling. Suppose, for instance, you are currently committed to saving neither John nor James from drowning; that's what you plan to do; it's the action you have formed an intention to perform. What's more, you're certain that this action is not moral. But you're uncertain which of the other two available actions is moral. And let's add a further twist to drive home the point. Suppose, furthermore, that you are certain that you are just about learn, of exactly one of them, that it is permissible. And add to that the fact that, immediately after you learn, of exactly one of them, that it is moral, you must act—failing to do so will leave both John and James to drown, and of course you are certain that this is immoral. In this case, I think, it's quite reasonable to say that, upon learning that saving James is permissible, you are not only morally permitted to drop your intention to save neither and replace it with the intention to save James, but you are also morally required to do so; and the same should you learn that it is permissible to save John. It would, I think, be impermissible to save neither, since you're certain that's immoral and you know of an alternative that is moral; and it would be impermissible to save John, since you are still uncertain about the moral status of that action, while you are certain that saving James is moral; and it would be morally required to save James, since you are certain of that action alone that it is moral.

Now, Levinstein's principle might seem to hold for individuals in an analogous situation. Suppose you're certain that your current credences are irrational. And suppose you're certain that you will learn of only one credence function that it is rationally permissible. At least in this situation, it might seem that it is rationally required that you adopt the credence function you learn is rationally permissible, just as you are morally required to perform the single act you learn is moral. So, is Levinstein's argument rehabilitated? After all, if you really are rationally required to adopt $C'(-|C' \in \mathbf{R}_E)$ in this situation should you learn $C' \in \mathbf{R}_E$, and if you are rationally required to adopt $C''(-|C'' \in \mathbf{R}_E)$ should you learn $C'' \in \mathbf{R}_E$, then you must have the conditional credences to which Levinstein appeals to derive his contradiction.

In fact, I don't think this rehabilitates Levinstein's argument. Thorstad's example is useful, but not because the case of rationality and morality are analogous; rather, the example is useful precisely because it draws attention to the fact that they are disanalogous. After all, all moral actions are better than all immoral ones. So, if you are committed to an action you know is immoral, and you learn of another that it is moral, and you know you'll learn nothing more about morality, you must commit to perform the action you've learned is moral. Doing so is the only way you know how to improve the action you'll perform for sure. But this is not the case for rational attitudes. It is not the case that all rational attitudes are better than all irrational attitudes. Let's see a few examples.

Suppose my preferences over a set of acts $a_1, \ldots, a_N$ are as follows, where N is some very large number:

$$a_1 \prec a_2 \prec a_3 \prec \ldots \prec a_{N-3} \prec a_{N-2} \prec a_{N-1} \prec a_N \prec a_{N-2}$$

These preferences are irrational, because, if the ordering is irreflexive, then it is not transitive: $a_{N-2} \prec a_{N-1} \prec a_N \prec a_{N-2}$, but $a_{N-2} \not\prec a_{N-2}$. Now, suppose I learn that the following preferences are rational:

$$a_1 \succ a_2 \succ a_3 \succ \ldots \succ a_{N-3} \succ a_{N-2} \succ a_{N-1} \succ a_N$$

Then surely it is not rationally required of me to adopt these alternative preferences. Indeed, it seems to me that it might even be rationally prohibited to transition from the first irrational set to the second rational set. But I don't need that stronger claim. In the end, my original preferences are irrational because of a small, localised flaw. But they nonetheless express coherent opinions about a lot of comparisons. And, concerning all of those comparisons, the alternative preferences take exactly the opposite view. Moving to the latter in order to avoid having preferences that are flawed in the way that the original set are flawed does not seem rationally required, and indeed might seem irrational.

Something similar happens in the credal case, at least according to the epistemic utility theorist. Suppose I have credence function C where $C(X) = 0.1$ and $C(\overline{X}) = 1$. And suppose the single legitimate measure of epistemic utility is the Brier score.[40] I don't know this, but I do know a few things: first, I know that the rationality of credences is determined by features of their epistemic utility in different possible worlds, and that whatever is the correct measure of epistemic utility it is truth-directed, so that higher credences in truths and lower credences in falsehoods are better; furthermore, I know that my credences are dominated and therefore irrational, but I don't know what dominates them. Now suppose I learn that the following credence function is rational: $C'(X) = 0.95$ and $C'(\overline{X}) = 0.05$. It seems that I am not required to adopt C'; and again, it seems that I am not even rationally permitted to do so, though again this latter claim is stronger than I need. While C is irrational, it does nonetheless encode something like a

[40]For arguments for this claim, see (D'Agostino & Sinigaglia, 2010) and (Pettigrew, 2016a, Section 4.4).

point of view. And, from that point of view, and relative to many many truth-directed measures of epistemic utility, the C' just looks much much worse than C. While I know that C is irrational and dominated, though I don't know what by, I also know that, from my current, slightly incoherent point of view, C' looks a lot less good than C. And indeed it will be much less good if X turns out to be false.

So, even in what seems to be the situation in which Levinstein's principle is most compelling, namely, when you are certain you're irrational and you will learn of only one credence function that it is rational, still it doesn't hold. It is possible to be sure that your credence function is an irrational response to your evidence, sure that an alternative is a rational response, and yet not be required to adopt the alternative because learning that the alternative is rational does not teach you that it's better than your current irrational credence function for sure—it might be much worse. This is different from the moral case. So, as stated, Levinstein's principle is false.

However, to make the deference argument work, Levinstein's principle need only hold in a single case. Levinstein describes a family of cases—those in which you're certain you're irrational—and claims that it holds in all of those. Thorstad's objection shows that it doesn't. I then narrowed the family of cases to avoid Thorstad's objection—perhaps Levinstein's principle holds when you're certain you're irrational and certain you'll only learn of one credence function that it's rational. After all, the analogous moral principle holds in those cases. But we've just seen that the doxastic version doesn't always hold there, because learning that an alternative credence function is rational does not teach you that it is better than your irrational credence function in the way that learning an act is moral teaches you that it's better than the immoral act you intend to perform. But perhaps we can narrow the range of cases yet further to find one in which the principle does hold.

Suppose, for instance, you are certain you're irrational, you know you'll learn of just one credence function that it's rational, and moreover you know you'll learn that it is better than yours. Thus, in the epistemic utility framework, suppose you'll learn that it dominates you. Then surely Levinstein's principle holds here? And this would be sufficient for Levinstein's argument, since, relative to any strictly proper measure of epistemic utility, each non-probabilistic credence function is dominated by many different probabilistic credence functions; so we could find the distinct C' and C'' we need to make Levinstein's reductio go through.

Not so fast, I think. How you should respond when you learn that C' is rational depends on what else you think about what determines the rationality of a credence function. Suppose, for instance, you think that a credence function is rational just in case it is not dominated, but you don't know which are the legitimate measures of epistemic utility. So, for instance, suppose you think there is only one legitimate measure of epistemic utility, and you know it's either the Brier score or the absolute value score.[41] And suppose your credence function is $C(X) = 0.1$ and $C(\overline{X}) = 1$, as above. Now you learn not only that $C^\star(X) = 0.05$ and $C^\star(\overline{X}) = 0.95$ is rational, but also that it dominates C and is not itself dominated. So you learn that $C^\star$ is epistemically better than C at all worlds, and there is no alternative that is better than $C^\star$ at all worlds. Then you thereby learn that the Brier score is the only legitimate measure of epistemic utility. After all, according to the absolute value score, $C^\star$ does not dominate C; in fact, C and $C^\star$ have exactly the same absolute value score at both worlds. You thereby learn that the credence functions that dominate C without themselves being dominated are those C' for which $C'(X)$ lies between the solution of $-(1-x)^2 - (1-x)^2 = -(1-0.05)^2 - (0-1)^2$ in $[0,1]$ and the solution of $-(0-x)^2 - (1-(1-x))^2 = -(0-0.05)^2 - (1-1)^2$ in $[0,1]$, and $C'(\overline{X}) = 1 - c(X)$. You are then permitted to pick any one of them—they are all guaranteed to be better than yours. You are not obliged to pick $C^\star$ itself.

The crucial point is this: learning that a credence function is rational teaches you something about the features of a credence function that determine whether it is rational. And that teaches you something about the set of rational credence functions—you learn it contains that credence function, but you might also learn other normative facts, such as the correct measure of epistemic utility, perhaps, or the correct decision rule to apply with the correct measure of epistemic utility to identify the rational credence functions. And learning those things may well require you to shift your current credences; but they need not compel you to adopt the credence function you initially learned was rational.

Indeed, you might be compelled to adopt something other than that credence. An example: suppose that, instead of learning that $C^\star$ is rational and dominates C, you learn that some other $C' \neq C^\star$ is rational and dominates C. Then, as before, you learn that the Brier score and not the absolute value score is the correct measure of epistemic utility, and thereby learn the set of credence functions that dominates yours. Perhaps rationality then requires you to fix up your credence function so that it is rational, but in a way that minimises the amount by which you change your current credences. How to measure this? Well, perhaps you're required to pick an undominated dominator C'' such that the expected epistemic utility of your current irrational credence function C from the point of view of this dominator C'' is maximal. That is, you pick the credence function that dominates you and isn't itself dominated and which thinks most highly of your original

[41] Recall:

- the Brier score: $\mathfrak{B}(C, w) = -\sum_{X \in \mathcal{F}} |w(X) - C(X)|^2$;
- the absolute value score: $\mathfrak{A}(C, w) = -\sum_{X \in \mathcal{F}} |w(X) - C(X)|$.

credence function. Measuring epistemic utility using the Brier score, this turns out to be the credence function $C^\star$ described above. Thus, given this reasonable account of how to respond when you learn what the rational credence functions are, upon learning that $C' \neq C^\star$ is rational, rationality then requires you to adopt $C^\star$.

In sum: For someone who is certain that their credence function is irrational, learning only that some alternative credence function is rational is not enough to compel them to move to that alternative; indeed, it's not enough to compel them to change their credences at all, since they've no guarantee that doing so will improve their situation. To compel them to change their credences, you must teach them how to improve their epistemic situation. But when you teach them that doing a particular thing will improve their epistemic situation, that usually teaches them normative facts of which they were uncertain before—how to measure epistemic value, or the principles for choosing credences once you've fixed how to measure epistemic value—and doing that will typically teach them other ways to improve their epistemic situation besides the one you've explicitly taught them. Sometimes there will be nothing to tell between all the ways they've learned to improve their epistemic situation, and so all will be permissible, as Thorstad imagines; and sometimes there will be reason to pick just one of those ways, and so that will be mandated, even if epistemic rationality is permissive. In either case, Levinstein's argument does not go through. The deference principle on which it is based is not true.

While Levinstein's deference principle avoids certain problems that Greco and Hedden's faces, Greco and Hedden do more to justify the sort of deference that both principles try to capture. Inspired by Edward Craig's (1999) account of the role of the concept of knowledge, and following the lead of Sinan Dogramaci (2012), Greco and Hedden argue for their deference principle by appealing to roles that they take the concept of rationality to play in our lives. They claim that we use this concept to identify individuals to whom to defer. Knowing that we must gain evidence about the world not only directly from our experience of it, but also from others, we need some concept to pick out those others from whom we are happy to learn. They claim this role is played by the concept of rationality. The deference principle follows almost analytically.

I have a couple of concerns about this argument. First, if we wish to pick out people in our community from whom we should be happy to learn, the concept of reliability serves our purposes better than the concept of rationality. After all, most philosophers agree that being rational requires more than being reliable; and yet reliability is all that really interests you if you wish to defer. So, if you use reliability instead of rationality to badge those to whom you should defer, you will learn more truths. So, if the concept of reliability already serves this purpose and indeed does it better than the concept of rationality, why would we have introduced the concept of rationality in the way that Dogramaci and Greco and Hedden claim we did?

Second, as Greco and Hedden concede, it seems that it is only the concept of epistemic rationality that was introduced for this purpose; the concept of practical rationality does not seem to demand deference, and therefore cannot have been introduced for the same purpose. If I know that you prefer the well-being of your child to the well-being of my niece, and I learn that you are rational, I am not obliged to defer to you on this. If I know that you prefer the life of the mind to the life of action, and I learn that you are rational, I am not obliged to defer to you on this. If I know that you weigh property rights more than equality in a political system, and I learn that you are rational, I am not obliged to defer to you on this. If I know that we would both obtain the same value from successfully climbing Annapurna I, and the same disvalue from an unsuccessful attempt, but I prefer not to climb it because I'm more risk-averse, and you prefer to climb it because you are more risk-inclined, and I learn that you are rational, I am not obliged to defer to you on this. Why, then, did we introduce a single concept—the concept of rationality—and then give a quite different account of it when we apply it to different attitudes?

Greco and Hedden consider this objection. They agree that practical rationality doesn't demand deference, and they consider what might lie behind the difference.

> [W]hile there is an objective, agent-neutral standard of correctness for beliefs (namely truth), there may not be an objective, agent-neutral standard of correctness for preferences. [...] it may be that preferences are correct only insofar as they line up with agent-relative facts about goodness, and this is why judging a preference to be rational doesn't involve a commitment to defer to it. (Greco & Hedden, 2016)

One problem with this is that we do not even defer to preferences of those rational individuals with whom we share the same agent-relative facts about goodness. That's what the mountain climbing example above illustrates. For you and I, the two outcomes—successfully climbing the mountain, and making an unsuccessful attempt—might have the same agent-relative goodness; we might value them both equally. But, because of our attitudes to risk, I prefer one action over another, while you prefer the second to the first. I might learn this and also learn that you are rational—after all, different attitudes to risk are rationally permitted. But, if I do, I am under no obligation to defer to you.

Of course, to some extent, this is why I reject deference in the case of epistemic rationality as well. Yes, you and I might both share the same epistemic goals—accuracy, perhaps, or accuracy and understanding together—and we might both measure how well we are achieving it using the same measure of epistemic value. But we might also have different attitudes to epistemic risk—after all, we follow William James in claiming that this is rationally permissible. If we do, and if I learn that you are rational, I have no obligation to defer to you. Indeed, I might well have an obligation not to defer to you, since your credences aren't derived by applying Bayes' Rule to priors that reflect my attitudes to risk.

7.2 The value question

Rationality is good; irrationality is bad. Most epistemologists would agree with this rather unnuanced take, regardless of their view of what exactly constitutes rationality and its complement. Granted this, a good test of a thesis in epistemology is whether it can explain why this is true. Can it answer the value question: Why is rationality valuable? And indeed Sophie Horowitz (2014) gives an extremely illuminating appraisal of different degrees of epistemic permissivism and impermissivism by asking of each what answer it might give. Her conclusion is that the extreme permissivist and the extreme impermissivist can give a satisfying answer to this question, or, at least, an answer that is satisfying from their own point of view; but the moderate permissivist cannot. In Horowitz's paper, the extreme permissivist is played by the extreme subjective Bayesian, who thinks that satisfying Probabilism and being certain of your evidence is necessary and sufficient for rational priors and updating by Bayes' Rule is necessary and sufficient to obtain rational posteriors; the extreme impermissivist is played by the objective Bayesian, who thinks that rationality requires the uniform prior and posteriors obtained from it by Bayes' Rule; and the moderate permissivist is played by the moderate subjective Bayesian, who thinks rationality imposes requirements more stringent than merely Probabilism, but who does not think they're stringent enough to pick out a unique credence function. In this section, I'd like to raise some problems for Horowitz's assessment, and try to offer my own answer to the value question on behalf of the moderate Bayesian. Of course, if we think that GHC is the strongest decision rule that applies to my priors, then extreme subjective Bayesianism is true, at least interpersonally. However, as I mentioned in Section 5.2, we might wish to strengthen GHC by ruling out some of the most extreme risk-inclined Hurwicz weights. And in this case, we'll obtain moderate subjective Bayesianism. So it behooves us to address Horowitz's worries.

Let's begin with the answers to the value question that Horowitz offers on behalf of the extreme permissivist and the impermissivist.

According to Horowitz, the extreme subjective Bayesian says this: (i) only by being rational can you have a credence function that is immodest, where a credence function is immodest if it uniquely maximises expected epistemic utility from its own point of view; and, (ii) when it comes to your doxastic attitudes, it's good for them to be immodest and bad for them to be modest.

On (i): Horowitz, like us, assumes that all legitimate epistemic utility functions are strictly proper, so that every probabilistic credence function expects itself to be better than any alternative credence function. From this, we can conclude that, on the extreme permissivist view—that is, the extreme subjective Bayesian view—rationality is sufficient for immodesty. It's trickier to show that rationality is also necessary for immodesty, because it isn't clear what we mean by the expected epistemic utility of a credence function from the point of view of a non-probabilistic credence function—the usual definitions of expectation make sense only for probabilistic credence functions. Fortunately, however, we don't have to clarify this much. We need only say that, at the very least, if one credence function is epistemically better than another at all possible worlds—that is, if the first dominates the second—then any credence function, probabilistic or not, will expect the first to be better than the second. We then combine this with Theorem 1(I) from above, which says that, if epistemic utility is measured by a stricty proper epistemic utility function, then each non-probabilistic credence function is dominated. This then shows that, for the extreme subjective Bayesian, being rational is necessary for being immodest.

On (ii): While modesty might be a virtue in the practical or moral sphere, it is a vice in the epistemic sphere. In the moral sphere, modesty, so long as it is not excessive, is a virtuous mode of self-doubt; it demands that you treat the achievements and moral judgments of others as worthy of respect. In the technical sense in which it is used here in the epistemic realm, modesty is irrational because it generates situations reminiscent of those paradoxical beliefs discussed by G. E. Moore, such as 'It's raining, but I don't believe it is'. Just as Moore's belief seems irrational, so it seems irrational both to believe something and to expect that some specific alternative opinion about it is epistemically better than yours. And similarly for credences. It seems irrational both to have a credence in a proposition and also to expect some specific alternative credence to be better than yours. If you really do expect that alternative to be better, why are you sticking with the credence you have? So, according to Horowitz's answer to the value question on behalf of the extreme permissivist, being rational is good and being irrational is bad because being rational is necessary and sufficient for being immodest; and it's good to be immodest and bad to be modest.

So much for the extreme permissivist. Let's turn now to the extreme impermissivist. Horowitz's answer to the value question on their behalf is much briefer: if you are rational, you maximise expected accuracy from the point of view of the one true rational credence function; if you are irrational, you don't; the former is good and the latter is bad.

Below, we'll return to the question of whether these answers are satisfying. But first I want to turn to Horowitz's claim that the moderate Bayesian cannot give a satisfactory answer. I'll argue that, if the two answers just given on behalf of the extreme permissivist and extreme impermissivist are satisfactory, there is a satisfactory answer that the moderate permissivist can give. Then I'll argue that, in fact, these answers aren't very satisfying.

Horowitz's strategy is to show that the moderate permissivist cannot find a good epistemic feature of credence functions that belongs to all that they count as rational, but does not belong to any they count as irrational. The extreme permissivist can point to immodesty; the extreme impermissivist can point to maximising expected epistemic utility from the point of view of the sole rational credence function. But, for the moderate, there's nothing—or so Horowitz argues.

For instance, Horowitz initially considers the suggestion that rational credence functions guarantee you a minimum amount of epistemic utility. As she notes, the problem with this is that either it leads to impermissivism, or it fails to include all and only the credence functions the moderate considers rational. Let's focus on the case in which we have opinions only about a proposition X and its negation $\overline{X}$—the point generalises. As usual, we represent a credence function C as a pair $(C(X), C(\overline{X}))$. And let's measure epistemic utility using the Brier score. So, when X is true, the epistemic utility of (p, q) is $-(1-p)^2 - q^2$, and when X is false, it is $-p^2 - (1-q)^2$. Then:

- For $r > -\frac{1}{2}$: there is no credence function that guarantees you at least epistemic value r. If you have at least that epistemic value at one world, you have less than that epistemic value at a different world.
- For $r = -\frac{1}{2}$, there is exactly one credence function that guarantees you at least epistemic value r. It is the uniform credence function $(0.5, 0.5)$.
- For $r < -\frac{1}{2}$, there are both probabilistic and non-probabilistic credence functions that guarantee you at least epistemic utility r.

So, Horowitz concludes, a certain level of guaranteed epistemic utility can't be what separates the rational from the irrational for the moderate permissivist, since for any level, either no credence function guarantees it, exactly one does, or there are both credence functions the moderate considers rational and credence functions they consider irrational that guarantee it.

She identifies a similar problem if we think not about guaranteed accuracy but about expected accuracy. Suppose, as the moderate permissivist urges, that some but not all probability functions are rationally permissible. Then, Horowitz claims, for many rational credence functions, there will be irrational ones that they expect to be better than they expect some rational credence functions to be. Horowitz gives the example of a case in which the rational credences in X are exactly those between 0.6 and 0.8 inclusive.

Then someone with credence 0.8 will expect the irrational credence 0.81 to be better than it expects the rational credence 0.7 to be—at least according to many many strictly proper measures of epistemic utility. So, Horowitz concludes, whatever separates the rational from the irrational, it cannot be considerations of expected epistemic utility.

I'd like to argue that, in fact, Horowitz should be happy with appeals to guaranteed or expected epistemic utility. Let's take guaranteed utility first. All that the moderate permissivist needs to say to answer the value question is that there are two things that rationality provides: immodesty and a guaranteed level of epistemic utility. Immodesty rules out all non-probabilistic credence functions, as we know from Theorem 1, while the guaranteed level of epistemic utility narrows further. How narrow depends on how much epistemic utility you wish to guarantee. So, for instance, suppose we say that the rational credence functions are exactly those $(p, 1-p)$ with $0.4 \leq p \leq 0.6$. Then each is immodest. And they are exactly the credence functions on X and $\overline{X}$ that have a guaranteed epistemic utility of at least $r = -(1-0.4)^2 - 0.6^2 = -0.72$. If Horowitz is satisfied with the immodesty answer to the value question when the extreme permissivist gives it, I think she should also be satisfied with it when the moderate permissivist combines it with a requirement not to risk certain low epistemic utilities (in this case, utilities below $r = -0.72$). And this combination of principles rules in all of the credence functions that the moderate counts as rational and rules out all they count as irrational.

Next, let's think about expected epistemic utility. Suppose that the set of prior credence functions that the moderate permissivist counts as rational is a closed convex set. For instance, perhaps the set of rational credence function is

$$\mathbf{R} = \{P = (p, 1-p) : 0.6 \leq p \leq 0.8\}$$

Then we can prove the following: if a credence function C is not in $\mathbf{R}$, then there is an alternative credence function $C^\star$ in $\mathbf{R}$ such that each P in $\mathbf{R}$ expects $C^\star$ to be better than it expects C to be.[42] Thus, just as Horowitz would have the extreme impermissivist answer the value question by saying that, if you're irrational, there's a credence function the unique rational credence function prefers to yours, while if you're rational, there isn't, the moderate permissivist can say that, if you're irrational, there is a credence function that all the rational credence functions prefer to yours, while if you're rational, there isn't.

Of course, you might think that it is still a problem for moderate permissivists that there are rational credence functions that expect some irrational credence functions to be better than some alternative rational ones. But

[42]For the proof strategy, see (Pettigrew, 2016a, Chapter 10), but replace the epistemically possible chance functions mentioned in that chapter with the rational credence functions in $\mathbf{R}$.

I don't think Horowitz will have this worry. After all, the same problem affects extreme permissivism, and she does not taken issue with this—at least, not in the paper we're considering. For any two probabilistic credence functions P_1 and P_2, there will be some non-probabilistic credence function P_1' that P_1 will expect to be better than it expects P_2 to be—P_1' is just a very slight perturbation of P_1 that makes it incoherent; a perturbation small enough to ensure it lies closer to P_1 than P_2 does.

A different worry about the account of the value of rationality that I have just offered on behalf of the moderate permissivist is that it seems to do no more than push the problem back a step. It says that all irrational credence functions have a flaw that all rational credence functions lack. The flaw is this: there is an alternative preferred by all rational credence functions. But to assume that this is indeed a flaw seems to presuppose that we should care how rational credence functions evaluate themselves and other credence functions. But isn't the reason for caring what they say exactly what we have been asking for? Isn't the person who posed the value question in the first place simply going to respond: OK, but what's so great about all the rational credence functions expecting something else to be better, when the question on the table is exactly why rational credence functions are so good?

This is a powerful objection, but note that it applies equally well to Horowitz's response to the value question on behalf of the impermissivist. There, she claims that what is good about being rational is that you thereby maximise expected accuracy from the point of view of the unique rational credence function. But without an account of what's so good about being rational, I think we equally lack an account of what's so good about maximising expected accuracy from the point of view of the rational credence functions.

So, in the end, I think Horowitz's answer to the value question on behalf of the impermissivist and my proposed expected epistemic utility answer on behalf of the moderate permissivist are unsatisfying.

What's more, Horowitz's answer on behalf of the extreme permissivist is also a little unsatisfying. The answer turns on the claim that immodesty is a virtue, together with the fact that precisely those credence functions identified as rational by subjective Bayesianism have that virtue. But is it a virtue? Just as arrogance in a person might seem excusable if they genuinely are very competent, but not if they are incompetent, so immodesty in a credence function only seems virtuous if the credence function itself is good. If the credence function is bad, then evaluating itself as uniquely the best seems just another vice to add to its collection.

So I think Horowitz's answer to the value question on behalf of the extreme permissivist is a little unsatisfactory. But it lies very close to an answer I find compelling. That answer appeals not to immodesty, but to non-dominance. Having a credence function that is dominated is bad. It leaves free epistemic utility on the table in just the same way that a dominated action in practical decision theory leaves free pragmatic utility on the table. For the extreme permissivist, what is valuable about rationality is that it ensures that you don't suffer from this flaw.

One noteworthy feature of this answer is the conception of rationality to which it appeals. On this conception, the value of rationality does not derive fundamentally from a positive feature, but from the lack of a negative feature. Ultimately, the primary notion here is irrationality. A credence function is irrational if it exhibits certain flaws, which are spelled out in terms of its success in the pursuit of epistemic utility. You are rational if you are free of these flaws. Thus, for the extreme permissivist, there is just one such flaw—being dominated. So the rational credences are simply those that lack that flaw—and the maths tells us that those are precisely the probabilistic credence functions.

We can retain this conception of rationality, motivate moderate permissivism, and answer the value question for it. In fact, there are at least two ways to do this. We have met something very close to one of these ways when we tried to rehabilitate the moderate permissivist's appeal to guaranteed epistemic utility above. There, we said that what makes rationality good is that it ensures that you are immodest and also ensures a certain guaranteed level of accuracy. But, a few paragraphs back, we argued that immodesty is not guaranteed to be a virtue. So that answer can't be quite right. But we can replace the appeal to immodesty with an appeal to non-dominance, and then the answer will be more satisfying. Thus, the moderate permissivist who says that the rational credence functions are exactly those $P = (p, 1-p)$ with $0.4 \leq p \leq 0.6$ can say that being rational is valuable for two reasons: (i) if you're rational, you aren't dominated; (ii) if you're rational you are guaranteed to have epistemic utility at least $r = -0.72$; (iii) only if you are rational will (i) and (ii) both hold. This answers the value question by appealing to how well credence functions promote epistemic utility, and it separates out the rational from the irrational precisely.

The other way to motivate moderate permissivism is to follow the strategy we've been developing throughout this book. The credence functions that are rational for you are those with maximal generalised Hurwicz scores relative to your generalised Hurwicz weights. If there are no restrictions on rational attitudes to epistemic risk, which we encode in our generalised Hurwicz weights, then we obtain extreme permissivism. For every probabilistic credence function C, there is some sequence Λ of Hurwicz weights such that C maximises $H^{\Lambda}(\mathfrak{EU}(-))$ for every strictly proper epistemic utility function $\mathfrak{EU}$. But, while we are permissive about attitudes to epistemic risk, we needn't think that anything goes. Perhaps there are rational restrictions on risk attitudes. Perhaps some risk-inclined attitudes are too extreme or beyond the pale. If so, then there will be probabilistic credence functions C for which there are no permissible Λ such that C maximises $H^{\Lambda}(\mathfrak{EU}(-))$. This gives moderate permissivism. And it simultaneously furnishes the moderate permissivist with an answer to the value question: if you are irrational, then there are no rational attitudes to epistemic risk that rationalise your credences; that is, your credences do not constitute a rational way to pursue the goal of epistemic value. The moderate permissivist can thereby answer the value question that Horowitz poses.

7.3 Living on the edge of rationality

In her 2018 paper, 'Living on the Edge', Ginger Schultheis issues a powerful challenge to epistemic permissivism about credences (Schultheis, 2018). The heart of the argument is the claim that a certain sort of situation is impossible. Schultheis thinks that all motivations for permissivism must render situations of this sort possible. Therefore, permissivism must be false, or at least these motivations for it must be wrong.

Here's the situation, where again we write $\mathbf{R}_E$ for the set of credence functions that it is rational to have when your total evidence is E.

- Our agent's total evidence is E;
- There is C in $\mathbf{R}_E$ and our agent knows that C is in $\mathbf{R}_E$;
- There is C' in $\mathbf{R}_E$ and our agent does not know that C' is in $\mathbf{R}_E$.

Schultheis claims that the permissivist must take this to be possible, whereas in fact it is impossible. Here are a couple of specific examples that the permissivist will typically take to be possible.

> **Example 1** You have credences only in X and $\overline{X}$; and you have evidence E. The credences it is rational for you to assign to X in response to E are precisely those in the closed interval $[0.4, 0.7]$. But you are not sure of the extent of this interval. For all you know, it might be any one of the following three intervals: $[0.41, 0.7]$ or $[0.39, 0.71]$ or $[0.4, 0.7]$. So you are sure that 0.5 is a rational credence in X, but you're not sure whether 0.4 is a rational credence in X. In this case:
>
> - $\mathbf{R}_E = \{P = (p, 1-p) : 0.4 \leq p \leq 0.7\}$
> - $C = (0.5, 0.5)$
> - $C' = (0.4, 0.6)$
>
> **Example 2** Suppose extreme subjective Bayesianism is true. So rationality requires only that your credences obey Probabilism; in particular, it does not require that they obey the Principle of Indifference. You know that they must obey Probablism, and

you know that it is rationally permissible to satisfy the Principle of Indifference, but you don't know whether or not rationality demands it. In this case:

- $\mathbf{R}_E$ is the set of probability functions;
- C is the uniform distribution required by the Principle of Indifference;
- C' is some probability function other than the uniform distribution.

Schultheis then appeals to a principle that she calls Weak Rationality Dominance. We say that one credence function C rationally dominates another C' if C is rational in all epistemic possibilities in which C' is rational, and also rational in some epistemic possibilities in which C' is not rational. Weak Rationality Dominance says that it is irrational to adopt a rationally dominated credence function. The important consequence of this for Schultheis' argument is that, if you know that C is rational, but you don't know whether C' is, then C' is irrational. As a result, in our example above, C' is not rational, contrary to what the permissivist claims, because it is rationally dominated by C. So permissivism must be false.

If Weak Rationality Dominance is correct, then, it follows that the permissivist must say that, for any body of evidence E and set $\mathbf{R}_E$ of rational responses, the agent with evidence E either must know of *every* credence function in $\mathbf{R}_E$ that it is in $\mathbf{R}_E$, or they must know of *no* credence function in $\mathbf{R}_E$ that it is in $\mathbf{R}_E$. If they know of some credence functions in $\mathbf{R}_E$ that they are in $\mathbf{R}_E$ and don't know of others in $\mathbf{R}_E$ that they are in $\mathbf{R}_E$, then the ones of whose membership of $\mathbf{R}_E$ they are ignorant are irrational according to Weak Rationality Dominance, and therefore they are not in $\mathbf{R}_E$, which gives a contradiction. But, whatever your reason for being a permissivist, it seems very likely that it will entail situations in which there are some credence functions that are rational responses to your evidence and that you know are such responses, while you are unsure about other credence functions that are, in fact, rational responses whether or not they are, in fact, rational responses. This is Schultheis' challenge.

I'd like to explore a response to Schultheis' argument that takes issue with Weak Rationality Dominance (WRD). I'll spell out the response in general to begin with, and then see how it plays out for the Jamesian version of permissivism I've been developing in this book.

One route into this response is it note that WRD seems to entail a deference principle of exactly the sort that I objected to in my discussion of Levinstein and Greco and Hedden above. Recall Levinstein's principle: if you are certain that you are irrational, and you learn that a credence function C' is rational, then you should adopt C'—or at least you should have the conditional credences that would lead you to do this if you were to apply Bayes' Rule. We might slightly strengthen Levinstein's version of the deference principle as follows: if you are unsure whether you are rational or not, and you learn that C' is rational, then you should adopt C'. WRD entails this deference principle. After all, suppose you have credence function C, and you are unsure whether or not it is rational. And suppose you learn that C' is rational (and don't thereby learn that C is also rational). Then, according to Schultheis' principle, you are irrational if you stick with C.

Above, I objected to Levinstein's deference principle, and others like it, because it relies on the assumption that all rational credence functions are better than all irrational credence functions. I think that assumption is false. I think there are certain sorts of flaw that render you irrational, and lacking those flaws renders you rational. But lacking those flaws doesn't ensure that you're going to be better than someone who has those flaws.

So I think that the rational deference principle is wrong, and therefore any version of WRD that entails it is also wrong. But perhaps there is a more restricted version of WRD that is both correct and capable of sinking permissivism about epistemic rationality. Consider, for instance, a restricted version of WRD that applies only to agents who have no credence function—that is, it applies to your initial choice of a credence function; it does not apply when you have a credence function and you are deciding whether to adopt a new one. This makes a difference. The problem with a version that applies when you already have a credence function C is this: even if C is irrational, it might nonetheless be better in some situation than the credence function C' that you learn to be rational; and it might be that C assigns a lot of credence to that situation. So it's hard to see how to motivate the move from C to C'. However, in a situation in which you have no credence function, and you are unsure whether C is rational (even though it is) and you're certain that C' is rational (and indeed it is), WRD's demand that you should not pick C seems more reasonable. You occupy no point of view such that C is less of a departure from that point of view than C' is. You know only that C' lacks the flaws for sure, whereas C might have them. Better, then, to go for C', is it not? And if it is, this is enough to defeat permissivism.

I think it's not quite that simple. I noted above that Levinstein's deference principle relies on the assumption that all rational credence functions are better than all irrational credence functions. Schultheis' WRD seems to rely on something even stronger, namely, the assumption that all rational credence functions are equally good in all situations. For suppose they are not. You might then be unsure whether C is rational (though it is) and sure that C' is rational (and it is), but nonetheless rationally opt for C because you know that C has some good feature that you know C' lacks and you're willing to take the risk of having an irrational credence function in order to open the possibility of having that good feature.

Here's an example, where we represent a credence function C on X and $\overline{X}$ as $(C(X), C(\overline{X}))$, as before. You are unsure whether $C = (0.7, 0.3)$ is rationally permissible. It turns out that it is, but you don't know that. On the other hand, you do know that $C' = (0.5, 0.5)$ is rationally permissible. But C and C' are not equally good in all situations. C' has the same accuracy whether X is true or false; in contrast, C is better than the first if X is true and worse than the first if X is false. C' does not open up the possibility of high accuracy that C does; though, to compensate, it does preclude the possibility of low accuracy, which the first doesn't. Surveying the situation, you think that you will take the risk. You'll adopt C, even though you aren't sure whether or not it is rational. And you'll do this because you want the possibility of being rational and having that higher accuracy. This seems a rationally permissible thing to do. So, it seems to me, WRD is false.

Although I think this objection is sufficient to show that WRD is false, I think it's helpful to see how WRD might play out for a particular motivation for permissivism. Instead of the Jamesian motivation for permissivism that I have been developing throughout the book, let's look at the other risk-sensitive motivation I outlined above while discussing Horowitz's value question. Some credence functions offer the promise of high epistemic utility—those that assign high credence to X and low credence to its negation will be very good, epistemically speaking, if X is true. However, they also open the possibility of great epistemic disutility—they will be very bad if X is false. Other credence functions neither offer great epistemic utility nor risk great epistemic disutility—the same middling credence in X and its negation guarantees the same medium epistemic utility whether or not X is true. Bearing that in mind, you might say that you are more risk-averse the higher is the minimum possible epistemic utility you are willing to tolerate. And you might set a threshold r, and say that the options that are rational for you are those undominated options whose minimum epistemic utility is at least r.

Now, suppose that you are in the sort of situation that Schultheis imagines. You are uncertain of the extent of the set $\mathbf{R}_E$ of rational responses to your evidence E. On the account we're considering, this must be because you are uncertain of your own attitudes to epistemic risk. Let's say that the threshold of minimum epistemic utility you're willing to tolerate is $r = -0.98$, but you aren't certain of that—you think r might be anything between -0.72 and -1.28. So you're sure that it's rational to assign anything between 0.4 and 0.6 to X, since doing that has a minimum epistemic utility above -0.72, but you're unsure whether it's rational to assign 0.7 to X, which has a minimum epistemic utility of -0.98. In this situation, is it rational to assign 0.7 to X? WRD thinks not; I think it is. Among the credence functions that you know for sure are rational, the ones that give you the highest possible epistemic utility are: (i) the one that assigns 0.4 to

X, and (ii) the one that assigns 0.6 to X. They have a minimum epistemic utility of -0.72, and they open up the possibility of an epistemic utility of -0.32, which is higher than the highest possible epistemic utility opened up by any others that you know to be rational. On the other hand, assigning 0.7 to X risks an epistemic utility of -0.98, but it also opens up the possibility of an epistemic utility of -0.18, which is considerably higher than is available from the credences you know to be rational. As a result, it doesn't seem irrational to assign 0.7 to X, even though you don't know whether it is rational from the point of view of your attitudes to risk, and you do know that assigning 0.6 is rational. This tells against WRD.

There is another possible response to Schultheis' challenge for those who like this sort of motivation for permissivism. You might simply say that, if your attitudes to risk are such that you will tolerate a minimum epistemic utility of r, then regardlesss of whether you know this fact, indeed regardless of your level of uncertainty about it, the rational credence functions are precisely those that have minimum epistemic utility at least r. This sort of approach is familiar from expected utility theory. Suppose I have credences in X and in $\overline{X}$. And suppose I face two options whose utility is determined by whether or not X is true or false. Then, regardless of what I believe about my credences in X and $\overline{X}$, I should choose whichever option maximises expected utility from the point of view of my actual credences. The point is this: if what it is rational for you to believe or to do is determined by some feature of you, whether it's your credences or your attitudes to risk, being uncertain about those features doesn't change what it is rational for you to do. This introduces a certain sort of externalism to our notion of rationality. There are features of ourselves—our credences or our attitudes to risk—that determine what it is rational for us to believe or do, which are nonetheless not luminous to us. But I think this is inevitable. Of course, we might might move up a level and create a version of expected utility theory that appeals not to our first-order credences but to our credences concerning those first-order credences—perhaps you use the higher-order credences to define a higher-order expected value for the first-order expected utilities, and you maximise that. But doing this simply pushes the problem back a step. For your higher-order credences are no more luminous than your first-order ones. And to stop the regress, you must fix some level at which the credences at that level simply determine the expectation that rationality requires you to maximise, and any uncertainty concerning those does not affect rationality. And the same goes in this case. So, given this particular motivation for permissivism, which appeals to your attitudes to epistemic risk in a slightly different way from the way in which our favoured Jamesian version does, it seems that there is another reason why WRD is false. If C is in $\mathbf{R}_E$, then it is rational for you, regardless of your epistemic attitude to its rationality.

7.4 The ills of brute shuffling

A final worry about permissivism is that it seems to render permissible a doxastic version of what Michael Bratman (2012) calls 'brute shuffling' in the context of permissivism about conative attitudes such as desires; or, at least, it seems unable to explain what is irrational about this behaviour. An initial pass at the idea is this. There is a restricted range of catalysts for rational belief change. New evidence is the obvious one—learning new evidence makes it rational to change your belief. But you might also reassess old evidence, choosing to weight it differently; or you might acquire a new concept, thereby enriching your conceptual scheme; or you might discover a logical truth that bears on the relationship between different pieces of evidence. But most people will agree that the list doesn't get much longer than this. And they will surely agree that it is irrational to change what you believe in the presence of no catalyst at all—no new evidence, no reassessment, no new concepts, etc. We might say that someone who does this indulges in brute shuffling of their beliefs: 'shuffling' because they change their beliefs; 'brute' because there is nothing that precipitates their doing so.

Now, suppose permissivism is true and there are three rationally permitted responses to our agent's current evidence—they are the credence functions C_1, C_2, and C_3. For instance, we might suppose that (i) the individual has no evidence at all, (ii) he assigns credences to three possible worlds $\mathcal{W} = \{w_1, w_2, w_3\}$, and (iii) he has attitudes to risk that are encoded in his generalised Hurwicz weights $\Lambda = (\frac{1}{3}, \frac{1}{2}, \frac{1}{6})$. Then, according to GHC, the credence functions that are rational for him are:

$$C_1 = \left(\frac{5}{12}, \frac{5}{12}, \frac{2}{12}\right) \quad C_2 = \left(\frac{5}{12}, \frac{2}{12}, \frac{5}{12}\right) \quad C_3 = \left(\frac{2}{12}, \frac{5}{12}, \frac{5}{12}\right).$$

Now suppose that he shuffles between these three credence functions in rapid succession—first C_1, then C_2, then C_3—while all the time gaining no evidence, gaining no new concepts, and engaging in no reassessment of his evidence, which is in any case empty. Then we would say that he is irrational. Yet the permissivist seems unable to say why, since they consider each of these a rational response to her evidence.

In fact, I think things aren't as bleak for the permissivist. There are, I think, two things they can say. First, they might refer back to the argument for Bayes' Rule we borrowed from Gallow in Section 6.3. When our agent moves from C_1 to C_2 without any new evidence, he does not maximise2 expected epistemic utility by the lights of C_1. And similarly when he moves from C_2 to C_3 without any new evidence. Throughout, his epistemic utility function remains the same because his evidence remains the same. Indeed, because his evidence is empty, his epistemic utility function is simply $\mathfrak{EU} = \mathfrak{EU}^k_\top$, for some strictly proper $\mathfrak{EU}$. So C_1 expects C_1 to be better that it expects C_2 to be. So he is irrational when he moves to C_2. And similarly, even if he were to move to C_2, then since C_2 expects itself to be best, it would then be irrational to move from C_2 to C_3.

However, I suspect this won't entirely satisfy the objector. They might concede that, for the person who still takes their existing credence function to be part of the authoritative basis for their decision making, it would be irrational to do something other than what maximises expected epistemic utility by its lights. But might he not simply stop taking it to be authoritative in this way? Might he not simply abandon the prior he picked at the beginning of his epistemic life—namely, C_1—look at the various priors permitted by his Hurwicz weights—namely C_1, C_2, C_3—and choose a different one—say, C_2—take that instead to be authoritative and choose by maximising expected epistemic utility with respect to that? And might he not, in doing so, do nothing irrational?

I think there is a sense in which he might do all of this while avoiding irrationality. It is the same sense in which I might stop taking my current desires—encoded in my current utility function—to be authoritative. Nearly everyone agrees that rationality permits many different conative states. You value the life of the mind more than a life of action; I value the latter more than the former. You value long, intricate, all-consuming life projects; I prefer to flit from one short, easily-completed project to another. You prefer a single romantic relationship at the centre of your life; I prefer a number of platonic relationships. In these cases, your preferences are rationally permissible, and so are mine. Now suppose that I start my life with a particular utility function. At a later point, I'm called upon to make a decision. There is a sense in which I should choose using that utility function. I'm irrational if I don't. But, were I to abandon it and pick another from among those that are rationally permissible, there seems to be a sense in which there would be nothing irrational about this.

How can we reconcile these two judgments? I think the following picture is right: The mental attitudes that form the basis of your decision making are (i) your doxastic attitudes, represented by a credence function, (ii) your conative attitudes, represented by a utility function, and (iii) your risk attitudes, represented by your generalised Hurwicz weights. At the beginning of your epistemic life, you have only the latter two—(ii) utilities, both epistemic and pragmatic, and (iii) generalised Hurwicz weights. You use those to pick your prior credences. And, from then on you have credences as well—that is, you have (i). If your total evidence is E, then your epistemic utility function is $\mathfrak{EU}^E_k$, for some strictly proper $\mathfrak{EU}$ and some k. Now, at every point after you have picked your priors, you face a choice. You can continue with the credences, utilities, and Hurwicz weights that you have at that point. And, if you do that, then you should choose by maximising expected epistemic utility at that point. But you can also choose to abandon any part of your basic mental furniture and start afresh with some new permissible piece instead. As we've seen, when it comes to choosing how to update your credences, if you retain your credences and change only your epistemic utility function, it won't make any difference to how you should update—for any strictly proper epistemic utility functions $\mathfrak{EU}$ and any evidence E, $\mathfrak{EU}^k_E$ will recommend Bayes' Rule. But you might instead choose to abandon your credences. And you might then choose them again in line with GHC applied to your Hurwicz weights, just as you did at the beginning of your epistemic life. Or you might be even more radical, and abandon not only your credences but also your Hurwicz weights, and then you might choose them afresh, and then use those to choose your new priors; and then you might update those new priors on your new evidence.

The choice whether to abandon credences, utilities, or Hurwicz weights, or to continue taking your current ones to be authoritative, is arational. You do not make it from a rational perspective, for there is no rational perspective from which you can; you are deciding whether or not to abandon the perspective from which you make decisions. But if you choose not to abandon it, you thereby commit to using your credences and utilities to make your next decision.

So there is a way of filling in the story of the brute shuffler on which he is irrational, and a way of filling it in on which he is not. If he does not choose to abandon C_1 as authoritative for his decision making, and nonetheless switches to C_2 without new evidence, he is irrational. If, on the other hand, he does choose to abandon C_1 and returns to the situation in which he uses GHC to pick priors, and then picks C_2, then he has not acted irrationally.

Although this account does allow for one sort of brute shuffler who is not irrational, I think we can explain why our intuitions suggest that they are always irrational. We might note, for instance, that, when we hear the description of his credal changes, we assume that he is not someone who abandons the authority of their existing credences every time. And that's because we rarely encounter people doing that in the real world. So, when we hear the story of the shuffler, we fill it in with the details that make the shuffling irrational, and thus judge it so.

Why does the real world furnish us with so few examples of people who abandon their beliefs, desires, or their attitudes to risk? Why do we so rarely encounter people who replace the very basis of their decision-making? There are two related reasons. First: Mostly, in our day to day existence, we live *inside* our beliefs, our desires, and our attitudes to risk. We inhabit them. And part of what that involves is taking them to be authoritative. We might question them, of course, and indeed we do every time new evidence comes in. But when we question them, we use them to evaluate themselves. When new evidence comes in—perhaps our total evidence is E before and E' after—and our epistemic utility function evolves from $\mathfrak{EU}_k^E$ to $\mathfrak{EU}_{k'}^{E'}$, our credences evaluate themselves as less than optimal, and they evaluate the posterior obtained via Bayes' Rule as the best alternative, so we abandon them in favour of that. And when no new evidence comes in, they evaluate themselves as well—but this time, since the epistemic utility function stays the same, they evaluate themselves as best.

This is a less extreme form of questioning than the one that might lead us to abandon them completely as the basis of our decision-making. We tend to engage in the latter when something prompts us to live, however briefly, outside our beliefs, taking a third-person perspective on our own mental attitudes—looking down on them from above, as it were, and realising something about them that prompts us to reappraise them. This can happen for a number of reasons. Major psychological crisis might be one. The realisation that the ways in which you've been thinking have led to serious mental ill health can lead you to reevaluate those ways of thinking. Another catalyst is the realisation that your attitudes—doxastic, conative, risk-oriented—are in some way arbitrary. You have them not because you carefully picked them out, but because of the environment in which you grew up and the influences exerted on you from outside sources. For many people, that can drain these attitudes of their authority. And it might even lead them to abandon those attitudes and set about trying to pick alternatives in an autonomous way, which they believe will endow them with attitudes that they will take to have authority. That is Kant's hope, at least for our conative attitudes. I think it is a vain hope myself. There is no fixed point from which to pick a new basis for your decision-making in a way that is more authentically yours and autonomously chosen than the one bestowed on you by your nature and environment. So, if this is your motivation for abandoning your credences, for instance, then I think you will be disappointed. But we needn't settle that vast dispute here. My point here is only that the sort of realisation that tends to motivate creatures like us to abandon the very basis of our approach to the world and replace it with something else happens only occasionally in our lives: only through psychological crisis or the vertiginous feeling that these central components of oneself are arbitrary artefacts of one's environment or something else equally dramatic.

The second reason that we don't often encounter people who abandon the basis of their decision-making is that to do so takes a great deal of effort, if it is possible to do intentionally at all. Because we are thinking about the toy case in which there are just three worlds, it can seem as if it would not be a major upheaval. But of course real people have credence functions over many thousands, possibly millions, of possible worlds. To shift all of these attitudes would be a major undertaking.

In sum, then, I understand the intuitive reaction to the brute shuffling case. It is an unusual case; the sort of thing we don't encounter much in the world. But, when we hear it, we hold fixed the fact that people rarely abandon fundamental parts of the basis for their decision-making; we hold fixed that, when the brute shuffler switches from C_1 to C_2, he is likely still taking C_1 to be authoritative, and similarly when he moves from C_2 to C_3. And if that is the case, he is indeed irrational, and the permissivist can say why.

For some impermissivists, even this error theory for our intuitive judgment about the brute shuffler won't suffice. For them, such judgments carry a lot of weight, and it disqualifies a view, or at least tells very strongly against it, if it cannot accommodate them. I find that strange. After all, consider how strong the impermissivist's claim is. For any body of evidence, they say, there is a unique credal response; among all of the possible credal responses that satisfy the minimum requirements for respecting that evidence—all those that assign it and all propositions that follow logically from it maximal credence—there is a single, privileged one that is the correct rational response. Due to the efforts of J. M. Keynes, Rudolf Carnap, E. T. Jaynes, Jon Williamson, Jeff Paris and Alena Vencovska, and others, we know how difficult it is to identify that privileged response and to argue that it is privileged (Keynes, 1921; Carnap, 1950; Jaynes, 2003; Williamson, 2010; Paris & Vencovská, 1990, 1997). Why think that our pretheoretical intuitions about rationality, which apparently entail this extremely strong position, carry any weight at all?

The situation seems to me analogous to David Hilbert's intuitive judgment—shared by many in the European mathematical community of the early 20th century—that there must be a complete axiomatization of mathematics. Perhaps that intuition was reasonable. But citing that intuition would have been a poor argument against a foundation for mathematics on which there was no such axiomatization—perhaps one on which there are many different mathematical universes in which different mathematical statements are true. Similarly, consider any intuitive judgment made prior to Kenneth Arrow's work in the middle of the twentieth century to the effect that there is a fully satisfactory and fully general voting system for a democracy. Again, I don't mean to deny that such an intuition would have been reasonable. But since it is an intuitive judgment concerning the existence of a non-trivial, theoretically complex entity, it's hard to see why the intuition itself should carry enough weight to tell against a philosophical position on which there is no such thing.

The same goes for permissivism. To point out that many of our intuitive judgments entail its negation tells us little more than that we indulge in wishful thinking in this area. We would like there to be a single rational response to each piece of evidence, but wanting it doesn't make it so. To defeat permissivism, our opponent must present a fully worked out account of the unique credal state demanded by each body of evidence. It cannot be refuted by noting that our hope that it is true has infected our intuitive judgments about rationality.

7.5 Priors that allow you to learn

We often learn from experience. Sometimes, the evidence we collect tells us something about the underlying mechanism that produced it, and that teaches us something about what that mechanism is likely to produce in the future. For instance, if I don't know the half-life of a sample of radioactive material, and I then observe its decay over a period, I learn something about its half-life, and that teaches me something about how I should expect it to decay in the future. Or if I don't know whether a particular person is a reliable source of information on a given topic, and I independently corroborate some of the things they've told me about it, I become more confident that they are reliable, and more confident in the other things they've told me. A significant concern about the account I've been sketching is that it permits attitudes to epistemic risk that demand priors that leave you unable to learn in such situations. I'll illustrate with a toy example, and explain what we might say in response.

Before you stands an urn. You know that it's filled with only green and purple balls. And you know that there are either exactly three purple balls and exactly one green ball ($U_{\frac{1}{4}}$), or there are exactly three green balls and exactly one purple one ($U_{\frac{3}{4}}$). You will now draw one ball and replace it: G_1 is the proposition that you draw a green ball on your first draw, P_1 is the proposition that you draw a purple one. Then you will draw another: G_2 is the proposition you draw green on the second draw, P_2 says you draw purple. Then the eight possible worlds are:

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8
$U_{\frac{1}{4}}G_1G_2$	$U_{\frac{1}{4}}G_1P_2$	$U_{\frac{1}{4}}P_1G_2$	$U_{\frac{1}{4}}P_1P_2$	$U_{\frac{3}{4}}G_1G_2$	$U_{\frac{3}{4}}G_1P_2$	$U_{\frac{3}{4}}P_1G_2$	$U_{\frac{3}{4}}P_1P_2$

Now suppose that you are risk-averse in a particular way—perhaps your Hurwicz weights are $\Lambda = \left(\frac{1}{128}, \frac{1}{128}, \frac{1}{64}, \frac{1}{32}, \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}\right)$—or even risk-neutral—perhaps they are $\Lambda = \left(\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}\right)$. Then GHC demands that your prior is the uniform distribution:

$$C^{\dagger} = \left(\frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}\right)$$

So your credence that the urn contains just one green ball is $\frac{1}{2}$, your credence that the first ball drawn will be green is $\frac{1}{2}$, and your credence that the second ball will be green is $\frac{1}{2}$. But now suppose that you make your first draw, and it's green—that is, you learn G_1. Then your new credence function is:

$$C^{\dagger}_{G_1} := \left(\frac{1}{4}, \frac{1}{4}, 0, 0, \frac{1}{4}, \frac{1}{4}, 0\right)$$

And so, while your new credence in G_1 is 1, your credence is G_2 is still $\frac{1}{2}$ (and the same for P_2) and your credence in $U_{\frac{1}{4}}$ is still $\frac{1}{2}$ (and the same for

$U_{\frac{3}{4}}$). So, while you have learned something from your first draw, namely, that G_1 is true and P_1 false, it has had no impact on your credences in $U_{\frac{1}{4}}$ and $U_{\frac{3}{4}}$, and no impact on your credences in G_2 and P_2. And yet we do think that it should be possible to learn about the composition of the urn from our first draw, and also about the outcome of the next draw from the urn. We think that learning that the first ball drawn was green should make us more confident that the urn contains three green balls and less confident that it contains only one, since the former makes our evidence more likely; and then, having become more confident that the urn contains three balls, we should be more confident that the second ball drawn will be green.

This is the well-known problem that Rudolf Carnap faced when he tried to formulate inductive logic by identifying unique rational priors—what we have called $C^\dagger$, the uniform distribution over the state descriptions of the worlds, Carnap called $\mathfrak{m}^\dagger$ (Carnap, 1950). The problem has been raised in the present context by Wayne Myrvold (2019).

The natural solution is to demand that our prior satisfies the Principal Principle. The idea is this: According to Bayes' Theorem and Bayes' Rule, if C is my prior, then my posterior credence in $U_{\frac{1}{4}}$, upon learning G_1, should be:

$$C_{G_1}(U_{\frac{1}{4}}) = C(U_{\frac{1}{4}}|G_1) = \frac{C(U_{\frac{1}{4}})C(G_1|U_{\frac{1}{4}})}{C(U_{\frac{1}{4}})C(G_1|U_{\frac{1}{4}}) + C(U_{\frac{3}{4}})C(G_1|U_{\frac{3}{4}})}$$

The problem with the uniform prior is that it makes G_1 independent of $U_{\frac{1}{4}}$ and independent of $U_{\frac{3}{4}}$. That is, $C^\dagger(G_1|U_{\frac{1}{4}}) = C^\dagger(G_1) = C^\dagger(G_1|U_{\frac{3}{4}})$. The Principal Principle, in contrast, says that, conditional on the urn containing one green ball and three purple ones, your prior credence that your first draw will be green should be $\frac{1}{4}$—that is, $C(G_i|U_{\frac{1}{4}}) = \frac{1}{4}$, for $i = 1, 2$—while conditional on the urn containing three green balls and one purple one, your prior credence that your first draw will be green should be $\frac{3}{4}$—that is, $C(G_i|U_{\frac{3}{4}}) = \frac{3}{4}$, for $i = 1, 2$. In other words, conditional on the chances being a particular way, your credences should match those chances.

Now, suppose we impose this constraint. Then, to determine all our credences in each world, we need only say how likely we consider each of the two distributions of balls within the urns. That is, we need only specify $C(U_{\frac{1}{4}})$ and $C(U_{\frac{3}{4}})$. After all, if $C(U_{\frac{1}{4}}) = \lambda$, then the Principal Principle demands the following credence function:[43]

$$C^\lambda := \left(\frac{\lambda}{16}, \frac{3\lambda}{16}, \frac{3\lambda}{16}, \frac{9\lambda}{16}, \frac{9(1-\lambda)}{16}, \frac{3(1-\lambda)}{16}, \frac{3(1-\lambda)}{16}, \frac{1-\lambda}{16}\right)$$

Thus, if I draw green on my first draw, then my posterior is:

$$C^\lambda_{G_1} = \left(\frac{\lambda}{12-8\lambda}, \frac{3\lambda}{12-8\lambda}, 0, 0, \frac{9(1-\lambda)}{12-8\lambda}, \frac{3(1-\lambda)}{12-8\lambda}, 0, 0\right)$$

So your posterior credence in $U_{\frac{1}{4}}$ is $\frac{\lambda}{12-8\lambda} + \frac{3\lambda}{12-8\lambda} = \frac{4\lambda}{12-8\lambda}$, which is less than your prior in $U_{\frac{1}{4}}$, which was λ. For instance, if $\lambda = \frac{1}{2}$, your prior in $U_{\frac{1}{4}}$ is $\frac{1}{2}$ and your posterior is $\frac{1}{4}$. And your credence in G_2 is $\frac{\lambda}{12-8\lambda} + \frac{9(1-\lambda)}{12-8\lambda} = \frac{9-8\lambda}{12-8\lambda}$, which is greater than your prior in G_2, which was $\frac{12-8\lambda}{16}$. For instance, if $\lambda = \frac{1}{2}$, then your prior in G_1 is $\frac{1}{2}$, while your posterior is $\frac{5}{8}$. So, as we would like, learning that the first draw is green makes you less confident that the urn contains only one green ball and more confident it contains three, and it makes you more confident that the second ball drawn will be green.

Now, the first thing to note is that GHC can accommodate this. That is, for any prior C^λ that satisfies the Principal Principle, there are Hurwicz weights such that GHC with those weights that permit C^λ. But these Hurwicz weights will be reasonably risk-inclined. For instance, suppose $C(U_{\frac{1}{4}}) = \lambda = \frac{1}{2}$ in our example. Then the prior that the Principal Principle requires is:

$$C^{\frac{1}{2}} = \left(\frac{1}{32}, \frac{3}{32}, \frac{3}{32}, \frac{9}{32}, \frac{9}{32}, \frac{3}{32}, \frac{3}{32}, \frac{1}{32}\right)$$

And if this prior is permitted by your Hurwicz weights $(\lambda_1, \ldots, \lambda_8)$, then $\lambda_1 + \lambda_2 \geq \frac{9}{16}$. That is, to borrow Myrvold's turn of phrase, satisfying the Principal Principle in this case is a risky business—it requires you to put the majority of your Hurwicz weight on the best and second-best case. And how risk-inclined the Principal Principle requires you to be becomes more extreme as the possible objective chances become more extreme. For instance, suppose we know instead that the urn contains either one green ball and seven purple ones or seven green ones and one purple. In that case, if I am indifferent between the two hypotheses about the urn's composition, and if I satisfy the Principal Principle, then my prior must be:

$$\left(\frac{1}{128}, \frac{7}{128}, \frac{7}{128}, \frac{49}{128}, \frac{49}{128}, \frac{7}{128}, \frac{7}{128}, \frac{1}{128}\right)$$

[43]For instance,

$$C^\lambda(U_{\frac{1}{4}}G_1P_2) = C^\lambda(U_{\frac{1}{4}})C^\lambda(G_1P_2|U_{\frac{1}{4}}) = \lambda\frac{1}{4}\frac{3}{4} = \frac{3\lambda}{16}$$

And that requires $\lambda_1 + \lambda_2 \geq \frac{49}{64}$. That's a very risk-inclined set of Hurwicz weights!

The above shows that GHC will permit both priors that prevent learning from experience (e.g. $C^\dagger$) and priors that will demand it (C^λ). But can we go further? Can we show that rationality requires us to learn from experience in the way that priors that satisfy the Principal Principle demand? As we saw in Section 3.3 above, there is an epistemic utility argument for the Principal Principle. Say that one credence function chance dominates another if each possible objective chance function expects the former to be better, epistemically speaking, than it expects the latter to be. Then the argument for the Principal Principle is based on the following fact: If your prior violates the Principal Principle, then there is an alternative that chance dominates it; what's more, there's an alternative that chance dominates it that satisfies the Principal Principle; and yet further, no credence function that satisfies the Principal Principle is chance dominated. Add to this mathematical fact the following decision rule, which we might call Undominated Chance Dominance: it is irrational to choose an option that is chance dominated by an alternative that is not itself chance dominated. And this gives an epistemic utility argument for the Principal Principle.

Take the uniform prior $C^\dagger$, for instance. Then, whichever strictly proper epistemic utility function you use, it turns out that one of the credence functions that each chance function—the one corresponding to $U_{\frac{1}{4}}$ and the one corresponding to $U_{\frac{3}{4}}$—expects to have higher epistemic utility is $C^{\frac{1}{2}}$, the prior we described above that satisfies the Principal Principle and assigns equal credence to the two hypotheses about the composition of the urn.[44]

So, applied to the choice of priors, Chance Dominance is more demanding than GHC coupled with permissivism about attitudes to epistemic risk. The latter permits the uniform distribution $C^\dagger$, for instance, together with every C^λ for $0 \leq \lambda \leq 1$; Chance Dominance permits every C^λ, but nothing more. How should we respond to this?

We might simply say that Chance Dominance imposes new rational constraints on the Hurwicz weights that it's rationally permissible for you to have. Only those that, when coupled with GHC, permit priors that satisfy the Principal Principle are permitted. As before, I'd prefer to avoid this sort of restriction if possible. Part of the reason in this case is that, as we saw above, which priors satisfy the Principal Principle depend on which objective chance functions are possible. If the chance of green on each draw is either $\frac{1}{4}$ or $\frac{3}{4}$, then equal credence in each possibility demands one prior;

[44]To calculate these expected values, each chance distribution must assign a chance to each world. That means that each must assign a chance to the propositions $U_{\frac{1}{4}}$ and $U_{\frac{3}{4}}$. We simply assume here that the function that gives the chances if $U_{\frac{1}{4}}$ is true assigns chance 1 to $U_{\frac{1}{4}}$ and 0 to $U_{\frac{3}{4}}$, while the function that gives the chances if $U_{\frac{3}{4}}$ is true assigns chance 1 to $U_{\frac{3}{4}}$ and 0 to $U_{\frac{1}{4}}$.

if the chance is either $\frac{1}{8}$ or $\frac{7}{8}$, then equal credence in each possibility demands another. But I don't think we can require your attitudes to risk to be sensitive to this.

Nonetheless, while I think Chance Dominance is not rationally compelling in this situation—I think we must admit the rationality of the inductive scepticism embodied in $C^{\dagger}$—I do think it is something to which we might appeal when we justify having prior C^{λ}, if that is indeed the prior we have. As noted, Chance Dominance can demand priors that are very risky, and only rationalised via GHC by picking very risk-inclined Hurwicz weights. And, as noted earlier in the book, we might be wary of the rationality of extremely risk-inclined weights. But Chance Dominance can go some way to justifying these: only by having these weights do we avoid credence functions that are chance dominated.

A second response points out that GHC is a decision rule designed for a situation in which probabilities are not available. In the situation we're considering, some probabilities are available: they are not our credences, of course, as we've still to pick those; but they are the possible objective chances. In such a situation, we might amend GHC. GHC says that you should prefer one option to another if the generalised Hurwicz score of the first exceeds the generalised Hurwicz score of the second. To calculate the generalised Hurwicz score of an option a, we order the worlds by the utilities that option obtains at them, assign those utilities weights according to their position in that order, and sum up the weighted utilities. So, if

$$a(w_{i_1}) \geq \ldots \geq a(w_{i_n})$$

and your Hurwicz weights are $\Lambda = (\lambda_1, \ldots, \lambda_n)$, then

$$H^{\Lambda}(a) := \lambda_1 a(w_{i_1}) + \ldots + \lambda_n a(w_{i_n})$$

And GHC says:

(i) your preference ordering should be: $a \preceq b$ iff $H^{\Lambda}(a) \leq H^{\Lambda}(b)$;

(ii) you should not choose option a if there is an alternative option a' such that $a \prec a'$.

Now, define the generalised chance Hurwicz score of an option a as follows. Suppose $ch_1, \ldots, ch_m$ is the set of possible objective chance functions. Then we order the chance functions by the expected utilities they assign to a, assign those expected utilities weights according to their position in that order, and sum up the weighted expected utilities. So, if

$$\sum_{w \in \mathcal{W}} ch_{i_1}(w) a(w) \geq \ldots \geq \sum_{w \in \mathcal{W}} ch_{i_m}(w) a(w)$$

and your chance Hurwicz weights are $\Gamma = (\gamma_1, \ldots, \gamma_m)$, then

$$CH^{\Gamma}(a) := \gamma_1 \sum_{w \in \mathcal{W}} ch_{i_1}(w)a(w) + \ldots + \gamma_m \sum_{w \in \mathcal{W}} ch_{i_m}(w)a(w)$$

And the generalised chance Hurwicz principle (GCHC) says:

(i) your preference ordering should be: $a \preceq b$ iff $CH^{\Lambda}(a) \leq CH^{\Lambda}(b)$;

(ii) you should not choose option a if there is an alternative option a' such that $a \prec a'$.

The attraction of this is that, just as no dominated option maximises $H^{\Lambda}(-)$, so no chance dominated option maximises $CH^{\Gamma}(-)$. And so, when we apply it to the choice of credences, Theorem 2 tells us that whatever maximises the generalised chance Hurwicz score must satisfy the Principal Principle.

Let's see what the generalised chance Hurwicz principle says about our toy example of the urn from above. The two possible chance functions are:

- ch_1, which is the chance function if $U_{\frac{1}{4}}$ is true and there is just one green ball, so that $ch_1(G_i) = \frac{1}{4}$;
- ch_2, which is the chance function if $U_{\frac{3}{4}}$ is true and there are three green balls, so that $ch_2(G_i) = \frac{3}{4}$.

Then suppose your chance Hurwicz weights are $\gamma_1 = \gamma$ and $\gamma_2 = 1 - \gamma$. Then:

- if you are risk-averse or risk-neutral, so that $\gamma \leq \frac{1}{2}$, then the credence function that maximises $CH^{\Gamma}(\mathfrak{EU}(-))$ is $C^{\frac{1}{2}}$;
- if you are risk-inclined, so that $\gamma > \frac{1}{2}$, then the credence function that maximises $CH^{\Gamma}(\mathfrak{EU}(-))$ is C^{γ}.

I'll be honest—I don't feel confident in adjudicating between GHC and GCHC. While I think it's reasonable to care about the objective expected epistemic utility of your prior, and while, if you do, it's clearly worrying to find out that there's an alternative that the objective chances expect to do better than yours for sure, I also think it's reasonable to care about the actual utility of your prior, and to be reasonably risk-averse about that. But, as we have seen, these can pull in different directions. So I think that we have to accept that it is permissible either to ignore the unequivocal demands of the possible objective chances, or to ignore the risk of actual low epistemic utility—that is, it is permissible either to use GHC or GCHC. If that's right, then inductive scepticism of the sort encoded in $C^{\dagger}$ is permissible, but so are priors like C^{λ} that permit learning in the ordinary, Bayesian way.

7.6 Clifford's shipowner, conspiracy theories, and choosing for others

Perhaps the most natural place to look for the final objection to our Jamesian approach to rational belief is in the essay to which James' 'The Will to Believe' is a response, namely, William Kingdon Clifford's 1877 paper, 'The Ethics of Belief' (Clifford, 1877 [1999]). And the most famous objection in that paper comes from Clifford's example of the shipowner. Here it is in Clifford's words:

> A shipowner was about to send to sea an emigrant-ship. He knew that she was old, and not overwell built at the first; that she had seen many seas and climes, and often had needed repairs. Doubts had been suggested to him that possibly she was not seaworthy. These doubts preyed upon his mind, and made him unhappy; he thought that perhaps he ought to have her thoroughly overhauled and refitted, even though this should put him at great expense. Before the ship sailed, however, he succeeded in overcoming these melancholy reflections. He said to himself that she had gone safely through so many voyages and weathered so many storms that it was idle to suppose she would not come safely home from this trip also. He would put his trust in Providence, which could hardly fail to protect all these unhappy families that were leaving their fatherland to seek for better times elsewhere. He would dismiss from his mind all ungenerous suspicions about the honesty of builders and contractors. In such ways he acquired a sincere and comfortable conviction that his vessel was thoroughly safe and seaworthy; he watched her departure with a light heart, and benevolent wishes for the success of the exiles in their strange new home that was to be; and he got his insurance-money when she went down in mid-ocean and told no tales.
>
> What shall we say of him? Surely this, that he was verily guilty of the death of those men. It is admitted that he did sincerely believe in the soundness of his ship; but the sincerity of his conviction can in no wise help him, because *he had no right to believe on such evidence as was before him*. He had acquired his belief not by honestly earning it in patient investigation, but by stifling his doubts. And although in the end he may have felt so sure about it that he could not think otherwise, yet inasmuch as he had knowingly and willingly worked himself into that frame of mind, he must be held responsible for it.

While Clifford's essay preceded James', we can nonetheless extract an objection to a Jamesian picture from Clifford's example. According to the

Jamesian, you might think, the shipowner's high credence in the seaworthiness of his vessel must be rationally permissible; and yet using that credence when he decided to put the ship to sea with so many people on board was immoral. And this violates a maxim that says: if you choose in a rationally permissible way using rationally permissible credences and morally permissible utilities, then your choice is morally permissible.

There are a number of things to clarify here before we can proceed.

First, it isn't clear from Clifford's example exactly how the shipowner came to have the high credence that his boat was seaworthy. Did he retain his prior and choose to ignore some of the evidence that he had so that, conditional on that impoverished evidence, his prior assigned high credence to the ship's safety, even though it would have assigned low credence to the ship's safety conditional on the total evidence? Or did he replace his original prior with a new prior and retain his total evidence so that, conditional on his total evidence, this new prior assigned high credence to its seaworthiness? According to the version of Jamesian epistemology that I have been developing in this book, only the latter is rationally permissible. A posterior obtained from your prior by conditioning only on some proper part of your evidence will not maximise expected epistemic utility as it is defined in Gallow's argument, which we presented in Section 6.3 above.

Second, even if we think the shipowner's high posterior credence in the ship's safety arose because he replaced his prior and updated the new one on his total evidence, the way Clifford tells the story, it's clear that the shipowner changes his mind not because of some Damascene conversion or moment of crisis, but because of a conscious effort on his part to manipulate his own opinion in a way that would more greatly profit him but would put many lives at risk. And, as we noted above, the Jamesian approach for which I have been arguing permits an arational choice to change your prior from one time t to the next t', but it does not permit you to choose to change it by a standard expected utility calculation in which you continue to consider your credences at t authoritative for your choice of credences at t'. After all, providing your current credence function is certain of your total evidence, it will always think that changing without new evidence is irrational. And yet that is what the shipowner does. Effectively, it is as if he takes a pill at t that will make him confident at t' in the ship's seaworthiness, and he chooses to do this at t when he has a low credence in the ship's seaworthiness—or certainly a low enough one that he should not be sending her to sea. So, again, if this is the case, then the Jamesian is not obliged to consider the shipowner's credences rational, and we can pinpoint his immorality to time t, when he has low credence in the ship's safety, but manipulates his future self into having higher credences, and indeed credences high enough that they will lead him to put the ship to sea.

So, in order to make Clifford's challenge to my version of Jamesian epistemology stick, we need a slightly different example. We need a case in which a risk-inclined individual assigns a high prior to something that then leads them to make a decision that exposes many others to a risk of harm, where they would not have made this decision had they had a more risk-averse, less extreme prior.

One place we might look for such examples is among conspiracy theorists. After all, as Rachel Fraser (2020) points out, Jamesian epistemology is a natural lens through which to understand some conspiracist thinking:

> If James is right that epistemic risk aversion is a matter of taste rather than of rational compulsion, then it is not obviously irrational to believe that the royals killed Diana on extremely scant evidence. If someone's worse fear is missing out on true beliefs, then urging them only to believe claims for which they have good evidence is pointless. Sure, if you only believe things for which you have good evidence, you are less likely to have false beliefs. But you will also miss out on a lot of true ones. In other words: for someone with a conspirator's attitude to epistemic risk, believing in line with the evidence would itself be irrational, just as it would be irrational to order vanilla ice-cream if what you really want is chocolate. Those who think conspiracy theorists are too credulous think conspiracy theorists overestimate the strength of evidence for their theories. But this need not be so: they might correctly assess that their evidence is weak but take it that even weak evidence makes belief a bet worth taking. (Fraser, 2020)

Thus, the conspiracist is a risk-inclined Jamesian believer with priors that assign a lot of credence to their favoured conspiracy theory. As is often noted, most conspiracy theories include not only an account of certain events or certain political structures that offers an alternative to the official or mainstream account, but also an account of how the conspiracy posited in this alternative account has been and will be covered up, and how the population has been and will be deceived about its existence and manipulated to ignore or dismiss it with misleading evidence. The result is that most publicly accessible evidence that non-conspiracists would take to confirm the official explanation in fact has no power to disconfirm the conspiracy theory for the conspiracist, since the conspiracy theory itself makes such evidence as likely as the official explanation does—"Well, isn't that exactly what you'd expect them to say if they were covering up Roswell?"; "Well, of course they'd choose an unassuming pizza parlour—what could be more natural?". So, just as their prior in the conspiracy theory is high, so is their posterior after updating that prior on the evidence. So, for instance, an anti-vaxxer might believe not only that Bill Gates has secreted nanotechnology inside the SARS-CoV-2 coronavirus vaccines that will help him manipulate those who receive them, but they will also believe that various government and medical officials are aware of this and are colluding with Gates, so that when we hear those officials pronounce in favour of the safety and efficacy of the vaccines, that evidence will do nothing or very little to lower the conspiracist's credence in Gates' malevolent plan—"Well, yes, that medical officer has proved trustworthy in the past, but that's because he's playing the long game; he's been getting us to trust him so he can pull this off".

So now let's consider such an individual. They begin with a high prior credence in their vaccination conspiracy, and they update that on their evidence, which does something to reduce their high prior but not much, since the theory to which they assign that prior makes likely exactly the evidence that non-conspiracists would find disconfirming. Now, of course, most, perhaps all, actual conspiracy theories that circulate in the real world are either internally inconsistent or strictly inconsistent with the evidence. But they need not be. It is perfectly possible to construct a consistent conspiracy theory that has the capacity to accommodate most apparently countervailing evidence. After all, we are used to positing something structurally analogous in the philosophy of science, where we take a successful theory and construct an alternative one that is empirically equivalent to the first, or at least accommodates all evidence gathered in favour of the original theory so far. Here, the official explanation corresponds to the original theory, and the conspiracist's alternative corresponds to the empirically equivalent alternative theory. Indeed, as Rachel Fraser points out, Descartes' sceptical hypothesis, which posits a malicious demon hellbent on deceiving me, is the original conspiracy theory in philosophy.

So, if our hypothetical anti-vaxxer assigns a high prior to her consistent but false conspiracy theory, and updates it on her evidence using Bayes' Rule, and ends up with a slightly lower but still high posterior, our Jamesian must consider her credences rationally permissible. Now suppose that she acts on those credences. Convinced that she is saving people from Gates' malicious control, she sabotages a vaccination centre, destroying their stock of the vaccine and delaying by a month the vaccinations of 1,000 elderly or other high risk people in the local area. Of these, fifty end up contracting COVID-19 and two die from it. I think many will say, as Clifford said of the shipowner, that this person's action was immoral, and she is morally responsible for the subsequent deaths.

So, again, we have a conflict with the maxim stated above, which says that actions chosen rationally on the basis of rationally permissible credences and morally permissible utilities issues are themselves morally permissible. How, then, might the Jamesian respond?

Here's the response I favour. We're used to thinking that, when we make choices that affect others, we're morally obliged to take into account the values and goals and desires of those others. Now, as well as taking into account the desires and values and other conative attitudes of the people whom your decision might affect, it seems to me that we should also take into account their attitude to risk.[45]

Hiking in the Scottish Highlands one day with my friend, the mist comes down.[46] To avoid getting separated, we rope ourselves to one another and agree I'll go ahead on the narrow path, taking the lead and making decisions about our route. At one point, I must choose whether to continue our ascent, giving us the chance of attaining the summit but risking very serious injury, or begin our descent, depriving us of the possibility we'll reach the top but removing any risk of harm. My friend and I have been hiking together for years, and I know we share exactly the same utilities for the possible outcomes—we value reaching the summit exactly as much, and similarly for avoiding injury; and we disvalue missing out on the summit exactly as much, and similarly for sustaining injury. But I also know that we have radically different attitudes to risk. While I am very risk-inclined, he is very risk-averse. Were he to face the choice instead of me, I know he'd choose to descend. It seems immoral, I think, for me to choose on behalf of the two of us to continue our attempt on the summit.

How morality requires us to take the risk attitudes of others into account is, of course, a thorny question, just as is the question how morality requires us to take the conative attitudes of others into account. There do seem to be some distinctive features of the case of risk.[47] Firstly, it seems that more risk-averse individuals carry greater weight.[48] Suppose that it was my friend who had taken the lead on the mountain path. When they reached the fork in the route, I suspect we would be happy to say that they'd be morally permitted to follow their own attitudes to risk and embark upon the descent. I doubt that the principle is as simple as this: choose as if you were the most risk-averse in the group of affected parties. After all, I think a single very risk-averse person in a population of risk-neutral people should not end up dictating all of the ethical choices that affect that population. But something in this vicinity might be correct.

In any case, if we are morally obliged to take into account the risk attitudes of others in practical matters when facing a choice that affects them, might we not also be obliged to take into account their attitudes to epistemic risk in setting the credences we use to make such decisions? And might this not account for the immorality of the anti-vaxxer's assault on the vaccination centre. If she knows that all the individuals affected by the decision shared her view, I think we'd be more inclined to see her action as morally permissible. The problem is that she did this based on priors that

[45]For other treatments of this claim, see (Buchak, 2017; Blessenohl, 2020; Rozen & Fiat, ms; Thoma, ms).

[46]For a fascinating philosophical discussion of rational attitudes to risk in the context of mountaineering, see (Ebert & Robertson, 2013).

[47]Blessenohl (2020) argues that there can be no satisfying general account of such choices.

[48]Rozen & Fiat (ms) offer an argument for certain instances of this general principle.

take significant epistemic risks; and these epistemic risks are unlikely to be tolerable to almost any of the people affected—the sort of risk-inclined attitude that leads the anti-vaxxer to assign her high prior in her conspiracy theory is pretty rare. So, however much she inhabits those beliefs, she must recognise that they are the product of particular attitudes to epistemic risk that few people share. And she is morally obliged to take that into account when she makes her decision whether or not to destroy the vaccine supply. So, according to this, our anti-vaxxer's credences may be rational, but acting on them in the way she does is immoral.

I think the foregoing explains why the Jamesian is right to think that it's possible to have rationally permissible credences on which it is immmoral to act. But there is another worry in the vicinity. According to a view that has been gaining considerable support recently, it is possible for an individual's belief to harm, and thus perhaps to be immoral, even if the individual does not act on it. For instance, according to a number of philosophers, a belief that a particular individual lives up to a negative stereotype about one of the demographic groups to which they belong might harm that person even if having the belief is based on sound statistical evidence (Basu, 2018; Basu & Schroeder, 2019; Basu, 2019). Suppose I know that, of the 200 academic staff in some part of my university, only one is a women; and I know that, of the 200 administrative staff, 150 are women. And now suppose that, when I visit that part of the university for the first time, I meet the single academic who is a woman in the lobby area as I enter. And suppose that, before she introduces herself, I form the belief that she is a member of the administrative staff. Then, according to these philosophers, I harm her. Similarly, then, we would expect a belief in a false conspiracy theory to harm members of any group, whether a whole race or ethnicity or gender or sexual orientation or religion, that it falsely maligns by wrongly attributing to them evil intentions, morally repugnant desires, and monstrous actions, even if the conspiracist believer were never to act on those beliefs. Antisemitic conspiracy theories furnish a frighteningly common example here.

Of course, as with the anti-vaxxers' conspiracies, antisemitic ones that actually circulate in the real world are typically internally inconsistent and inconsistent with accepted evidence. But, as before, it is possible to formulate versions that are internally consistent and that contain sufficient internal protection against existing evidence that non-conspiracists would count as strongly disconfirming. So the Jamesian must grapple with the fact that a high prior credence in such a theory is rationally permissible on their account, as is a slightly lower but still high posterior credence—again, slightly lower because presumably the protection against falsifying evidence is not total; some of the countervailing evidence has some effect on reducing the prior credence, but not much. But if this is the case, it is rationally permissible to believe some of the vilest, most toxic conspiracy theories that people have constructed. And if, as Basu and others argue, such beliefs can harm whether or not they issue in harmful actions, then it is possible to have epistemically rationally permissible credences that are extremely harmful.

In the end, I think the Jamesian must accept this. But it is open to them to say that, epistemically speaking, these beliefs are rationally permissible, while morally speaking, they are despicable and repugnant. And it is then open to them to say that, all things considered, we should not have these beliefs, because the level of moral depravity outweighs the rational permissibility.

In brief...

In this final chapter, I considered five objections to permissivism about epistemic rationality.

The first, given voice by Daniel Greco and Brian Hedden (2016) and also Ben Levinstein (2017), begins with the claim that we should defer to rationality. That is, at least if we aren't sure whether our own current credence function is rational, or if we are certain that it is not, then if we learn of an alternative credence function that it is rational, we should adopt it as our own. The objection then proceeds to show that, if permissivism is true, it is impossible to satisfy this norm of deference coherently. I respond by denying the norm of deference. For the epistemic utility theorist, if you are ignorant of some fact about rationality—such as the fact that a particular credal state is rationally permissible—then you must be ignorant of an axiological fact about how to measure epistemic utility or a normative fact about how to pick credal states on the basis of their epistemic utility. Learning that a particular credal state is rational thus teaches you something about one or other or both of these, and therefore something about what is rationally permitted. But this will typically not mandate that you must adopt as your credal state the one that you learned was rational. It might permit a number of different responses; or it might mandate that you respond by adopting a different credal state that your newly-acquired evidence about the grounds of rationality has also revealed to be rational.

The second objection is due to Sophie Horowitz (2014) and challenges the permissivist to say what is valuable about rationality. I respond that the epistemic utility theorist is particularly well placed to answer this question. For them, credences are irrational if they are flawed in certain ways as means by which to pursue the goal of epistemic value. They are rational if they lack these flaws. It is clear, then, why they are valuable—epistemically valuable credences are, by their nature, valuable; and rational credences are appropriate means by which to pursue epistemic value.

The third objection comes from Ginger Schultheis (2018), who begins by noting that nearly all permissivists will think there are situations in which you are partially ignorant of which credal states are rationally permissible, knowing of one that is rationally permissible that it is, and not knowing of another that is rationally permissible that it is. But, Schultheis argues, such a situation is impossible. She contends that it is never rational to adopt a state that you don't know is rational when there is another that you do know to be rational—she calls this principle Weak Rational Dominance. I respond by objecting to this principle. I note that not all rational credal states are equally good in all situations—they exhibit different virtues in different circumstances. I argue that even if I am not sure whether one state is rational, I might rationally opt for it instead of one I know to be rational because in some circumstances it has a virtue that the one I know to be rational always lacks, and I'm willing to take the risk of irrationality to secure the possibility of that virtue.

The fourth objection, offered by Wayne Myrvold (2019) to my earlier work on epistemic risk, notes that our account permits priors, such as the uniform prior, that do not allow you to learn from experience in the ways we think you should. In the end, I think these priors, and the inductive scepticism they embody, must be permitted, however little we like the consequences of adopting them. But, unlike in my earlier work, they are not now mandated.

The fifth and final objection I consider comes from W. K. Clifford's essay, 'The Ethics of Belief' (1877 [1999]). Since our account results in such a wide interpersonal permissivism, an individual with very risk-inclined Hurwicz weights will be permitted to assign so high a prior credence to a conspiracy theory that, even after rationally incorporating the countervailing evidence, they will end up with a high rational posterior in that theory. And yet it seems immoral for them to act on that credence. In our example, it seems immoral for an anti-vaxxer with high credence that a particular vaccine is harmful to destroy stocks of it, even though that action is expected to do the most good from the point of view of their credence. Does this cast doubt on the rationality of this credence? I argue that it does not. In the practical case, we are happy to say that, when I make a decision that will affect other people—such as the decision to destroy vaccine stocks—morality requires me to take into account their attitudes to pragmatic risk. I argue that the same holds true of their attitudes to epistemic risk. It is for this reason that the anti-vaxxer is morally required not to destroy the vaccine stocks—if they do so, they impose their own attitudes to epistemic risk on the people whose lives their decision will affect; and while that doesn't affect the rationality of their credences, it does affect the morality of acting on them.

8 Summing up

According to epistemic utility theory, the rational credal states are those it would be rational for you to pick were you presented with a choice between all the possible such states. To make such a choice, you'd need: (i) a way of measuring the purely epistemic value of each credal state at each way the world might be, and (ii) a decision rule that separates out the rational choices from the irrational ones on the basis of that epistemic value. To epistemic utility theory, we added William James' permissivism about attitudes to epistemic risk. Noting that we could not obtain permissivism by encoding our attitudes to epistemic risk in our measures of epistemic value, we turned instead to the decision rule we use to pick our credal state. We argued that, when we pick our prior credences, we should use a decision rule that encodes our attitudes to epistemic risk. In particular, it should be a decision rule from the family GHC: that is, we should represent our attitudes to risk by generalised Hurwicz weights and then pick a prior that has maximal generalised Hurwicz score relative to those weights. This gives us wide interpersonal and narrower intrapersonal permissivism about prior credences. When we come to pick our posteriors, on the other hand, we should use our priors along with the risk-neutral expected utility rule and an epistemic utility function suited to our evidence. After all, we've already encoded our attitudes to risk in the decision rule we used to pick our priors, and we're appealing to those priors to make this choice. This leads us to update our priors on our evidence using Bayes' Rule to give our posteriors. And this gives wide interpersonal and narrow intrapersonal permissivism about posteriors.

So where do we go from here? As always, there's much still to do.

(1) We might explore how to extend the framework described here:

  (a) We might ask how the approach works in the infinite case, where we lift our restriction to finite sets of possible worlds.

  (b) We might ask what the consequences of the approach are when we apply it not to credences but to beliefs, or other doxastic states.

(2) We might ask whether the way I have incorporated our attitudes to epistemic risk into the account is the only way this might be done. You might think, for instance, that you could incorporate some portion of our attitudes to risk into the decision rule we use to pick our priors and some portion into the rule we use to pick our posteriors. What different verdicts about permissivism would that give?[49]

[49]For important work in this direction, see (Campbell-Moore & Salow, 2020, ta).

(3) I've found sadly little to say about which attitudes to risk we should consider rationally permissible. We can be very permissive about such attitudes without countenancing all of them. But where, if anywhere, to draw the line? I feel confident, though, that wherever that line lies, it will include within it sufficiently many permissible attitudes to risk that they will entail permissivism about credal rationality.

(4) I have also said almost nothing about the implications for social epistemology, except in response to the Cliffordian objection above, where I suggested that the epistemically risk-inclined should not use their extreme priors when making decisions that will affect a group of individuals who are epistemically risk-averse.

    There are many questions in social epistemology on which the ideas in essay bear. I'll mention only two from network epistemology:

    (a) What are the advantages and disadvantages of different distributions of attitudes to epistemic risk throughout a population? Is it good to include some risk-inclined individuals in your epistemic group, even if you are yourself epistemically risk-averse? And does this provide an answer to Clifford's conservatism?[50]

    (b) While I have focused on individuals who have attitudes to epistemic risk, perhaps groups do as well, and perhaps risk-inclined groups will prefer epistemic networks with different structures from those that risk-averse groups will choose.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica*, *21*(4), 503–546.

Arló-Costa, H., & Pedersen, A. P. (2012). Belief and Probability: A General Theory of Probability Cores. *International Journal of Approximate Reasoning*, *53*(3), 293–315.

Basu, R. (2018). Can Beliefs Wrong? *Philosophical Topics*, *46*(1), 1–17.

Basu, R. (2019). The Wrongs of Racist Beliefs. *Philosophical Studies*, *176*(9), 2497–2515.

Basu, R., & Schroeder, M. (2019). Doxastic Wronging. In B. Kim, & M. McGrath (Eds.) *Pragmatic Encroachment in Epistemology*. Routledge.

[50]For important work in this direction, see (Zollman, 2010).

Berker, S. (2013). Epistemic Teleology and the Separateness of Propositions. *Philosophical Review*, *122*(3), 337–393.

Bernoulli, D. (1738 [1954]). Exposition of a New Theory on the Measurement of Risk. *Econometrica*, *22*(1), 23–36.

Blackwell, K., & Drucker, D. (2019). When propriety is improper. *Philosophical Studies*, *176*(2), 367–386.

Blessenohl, S. (2020). Risk Attitudes and Social Choice. *Ethics*, *130*, 485–513.

Bradley, S. (2016). Imprecise Probabilities. In E. N. Zalta (Ed.) *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Bratman, M. (2012). Time, Rationality, and Self-Governance. *Philosophical Issues (Supp. Noûs)*, *22*(1), 73–88.

Brier, G. W. (1950). Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review*, *78*, 1–3.

Briggs, R. A. (2009). Distorted Reflection. *Philosophical Review*, *118*(1), 59–85.

Briggs, R. A., & Pettigrew, R. (2020). An accuracy-dominance argument for conditionalization. *Noûs*, *54*(1), 162–181.

Buchak, L. (2013). *Risk and Rationality*. Oxford, UK: Oxford University Press.

Buchak, L. (2014). Belief, credence, and norms. *Philosophical Studies*, *169*, 285–311.

Buchak, L. (2017). Taking Risks Behind the Veil of Ignorance. *Ethics*, *127*(3), 610–644.

Campbell-Moore, C., & Salow, B. (2020). Avoiding Risk and Avoiding Evidence. *Australasian Journal of Philosophy*, *98*(3), 495–515.

Campbell-Moore, C., & Salow, B. (ta). Accurate Updating for the Risk-Sensitive. *British Journal for the Philosophy of Science*.

Carnap, R. (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Clifford, W. K. (1877 [1999]). The Ethics of Belief. In T. Madigan (Ed.) *The ethics of belief and other essays*. Amherst, MA: Prometheus.

Craig, E. (1999). *Knowledge and the State of Nature*. Oxford University Press.

D'Agostino, M., & Sinigaglia, C. (2010). Epistemic Accuracy and Subjective Probability. In M. Suárez, M. Dorato, & M. Rédei (Eds.) *EPSA Epistemology and Methodology of Science: Launch of the European Philosophy of Science Association*, (pp. 95–105). Springer Netherlands.

De Bona, G., & Staffel, J. (2017). Graded Incoherence for Accuracy-Firsters. *Philosophy of Science*, *84*(2), 189–213.

Dogramaci, S. (2012). Reverse Engineering Epistemic Evaluations. *Philosophy and Phenomenological Research*, *84*, 513–530.

Dorst, K. (2019). Lockeans Maximize Expected Accuracy. *Mind*, *128*(509), 175–211.

Easwaran, K. (2013). Expected Accuracy Supports Conditionalization - and Conglomerability and Reflection. *Philosophy of Science*, *80*(1), 119–142.

Easwaran, K. (2014). Decision Theory without Representation Theorems. *Philosophers' Imprint*, *14*(27), 1–30.

Easwaran, K. (2016). Dr Truthlove, Or: How I Learned to Stop Worrying and Love Bayesian Probabilities. *Noûs*, *50*(4), 816–853.

Ebert, P. A., & Robertson, S. (2013). A Plea for Risk. *Royal Institute of Philosophy Supplement*, *73*, 45–64.

Elga, A. (2010). Subjective Probabilities Should Be Sharp. *Philosophers' Imprint*, *10*(5), 1–11.

Foley, R. (1992). The Epistemology of Belief and the Epistemology of Degrees of Belief. *American Philosophical Quarterly*, *29*, 111–124.

Fraser, R. (2020). Epistemic FOMO. *The Cambridge Humanities Review*, *16*, 1–1.

Gallow, J. D. (2019). Learning and Value Change. *Philosophers' Imprint*, *19*, 1–22.

Greaves, H., & Wallace, D. (2006). Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind*, *115*(459), 607–632.

Greco, D., & Hedden, B. (2016). Uniqueness and Metaepistemology. *Journal of Philosophy*, *113*(8), 365–395.

Gustafsson, J. E. (ms). Decisions under Ignorance and the Individuation of States of Nature. Unpublished manuscript.

Hájek, A. (2008). Arguments For—Or Against—Probabilism? *The British Journal for the Philosophy of Science*, *59*(4), 793–819.

Hall, N. (1994). Correcting the Guide to Objective Chance. *Mind*, *103*, 505–518.

Hempel, C. (1962). Deductive-Nomological vs. Statistical Explanation. In H. Feigl, & G. Maxwell (Eds.) *Minnesota Studies in the Philosophy of Science*, vol. III, (pp. 98–169). Minneapolis: University of Minnesota Press.

Horowitz, S. (2014). Immoderately rational. *Philosophical Studies*, *167*, 41–56.

Horowitz, S. (2017). Epistemic Value and the Jamesian Goals. In J. Dunn, & K. Ahlstrom-Vij (Eds.) *Epistemic Consequentialism*. Oxford University Press.

Huber, F. (2006). The Consistency Argument for Ranking Functions. *Studia Logica*, *82*, 1–28.

Hurwicz, L. (1951). A Class of Criteria for Decision-Making Under Ignorance. Tech. Rep. Statistics 356, Cowles Commission Discussion Paper.

Hurwicz, L. (1952). A criterion for decision making under uncertainty. Tech. Rep. Technical Report 355, Cowles Commission.

Jackson, E. (2019). A Defense of Intrapersonal Belief Permissivism. *Episteme*, (pp. 1–15).

James, W. (1897). The Will to Believe. In *The Will to Believe, and other essays in popular philosophy*. New York: Longmans Green.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science.*. Cambridge, UK: Cambridge University Press.

Joyce, J. M. (1998). A Nonpragmatic Vindication of Probabilism. *Philosophy of Science*, *65*(4), 575–603.

Joyce, J. M. (2009). Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief. In F. Huber, & C. Schmidt-Petri (Eds.) *Degrees of Belief*. Springer.

Kelly, T. (2014). Evidence Can Be Permissive. In M. Steup, J. Turri, & E. Sosa (Eds.) *Contemporary Debates in Epistemology*, (pp. 298–311). Wiley-Blackwell.

Keynes, J. M. (1921). *A Treatise on Probability*. London: Macmillan.

Konek, J. (2016). Probabilistic Knowledge and Cognitive Ability. *Philosophical Review*, *125*(4), 509–587.

Konek, J. (2019). Comparative Probabilities. In R. Pettigrew, & J. Weisberg (Eds.) *Open Handbook of Formal Epistemology*, (pp. 267–348). PhilPapers Foundation.

Konek, J. (ta). Epistemic Conservativity and Imprecise Credence. *Philosophy and Phenomenological Research*.

Kyburg, H. E. (1961). *Probability and the Logic of Rational Belief*. Middletown: Wesleyan University Press.

Leitgeb, H. (2014). The Stability Theory of Belief. *Philosophical Review*, *123*(2), 131–171.

Leitgeb, H., & Pettigrew, R. (2010a). An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science*, *77*, 201–235.

Leitgeb, H., & Pettigrew, R. (2010b). An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy. *Philosophy of Science*, *77*, 236–272.

Levinstein, B. A. (2012). Leitgeb and Pettigrew on Accuracy and Updating. *Philosophy of Science*, *79*(3), 413–424.

Levinstein, B. A. (2017). Permissive Rationality and Sensitivity. *2017*, *94*(2), 342–370.

Lewis, D. (1980). A Subjectivist's Guide to Objective Chance. In R. C. Jeffrey (Ed.) *Studies in Inductive Logic and Probability*, vol. II. Berkeley: University of California Press.

Lewis, D. (1994). Humean Supervenience Debugged. *Mind*, *103*, 473–490.

Lewis, D. (1999). Why Conditionalize? In *Papers in Metaphysics and Epistemology*, (pp. 403–407). Cambridge, UK: Cambridge University Press.

Milnor, J. (1954). Games against nature. In D. T. R. Coombs, C. Davis (Ed.) *Decision Processes*, (pp. 49–60). New York: Wiley.

Myrvold, W. C. (2019). Learning is a Risky Business. *Erkenntnis*, *84*, 577–584.

Oddie, G. (1997). Conditionalization, Cogency, and Cognitive Value. *British Journal for the Philosophy of Science*, *48*, 533–41.

Paris, J. B., & Vencovská, A. (1990). A note on the inevitability of maximum entropy. *International Journal of Approximate Reasoning*, *4*, 181–223.

Paris, J. B., & Vencovská, A. (1997). In Defense of the Maximum Entropy Inference Process. *International Journal of Approximate Reasoning*, *17*, 77–103.

Peterson, M. (2004). From Outcomes to Acts: A Non-Standard Axiomatization of the Expected Utility Principle. *Journal of Philosophical Logic*, *33*, 361–378.

Pettigrew, R. (2013). A New Epistemic Utility Argument for the Principal Principle. *Episteme*, *10*(1), 19–35.

Pettigrew, R. (2016a). *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.

Pettigrew, R. (2016b). Accuracy, Risk, and the Principle of Indifference. *Philosophy and Phenomenological Research*, *92*(1), 35–59.

Pettigrew, R. (2016c). Jamesian epistemology formalised: an explication of 'The Will to Believe'. *Episteme*, *13*(3), 253–268.

Pettigrew, R. (2019). *Choosing for Changing Selves*. Oxford, UK: Oxford University Press.

Pettigrew, R. (2020). *The Dutch Book Argument*. Elements in Decision Theory and Philosophy. Cambridge, UK: Cambridge University Press.

Predd, J., Seiringer, R., Lieb, E. H., Osherson, D., Poor, V., & Kulkarni, S. (2009). Probabilistic Coherence and Proper Scoring Rules. *IEEE Transactions of Information Theory*, *55*(10), 4786–4792.

Quiggin, J. (1982). A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization*, *3*, 323–43.

Quiggin, J. (1993). *Generalized Expected Utility Theory: The Rank-Dependent Model*. Kluwer Academic Publishers.

Rozen, I. N., & Fiat, J. (ms). Attitudes to Risk when Choosing for Others. Unpublished manuscript.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons.

Savage, L. J. (1971). Elicitation of Personal Probabilities and Expectations. *Journal of the American Statistical Association*, *66*(336), 783–801.

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

Schoenfield, M. (2016). An Accuracy-Based Approach to Higher-Order Evidence. *Philosophy and Phenomenological Research*, *96*(3), 690–715.

Schoenfield, M. (2017). Conditionalization does not (in general) Maximize Expected Accuracy. *Mind*, *126*(504), 1155–87.

Schoenfield, M. (2019). Permissivism and the Value of Rationality: A Challenge to the Uniqueness Thesis. *Philosophy and Phenomenological Research*, *99*(2), 286–297.

Schultheis, G. (2018). Living on the Edge: Against Epistemic Permissivism. *Mind*, *127*(507), 863–879.

Staffel, J. (2016). Beliefs, Buses and Lotteries: Why belief can't be stably high credence. *Philosophical Studies*, *173*, 1721–1734.

Thau, M. (1994). Undermining and Admissibility. *Mind*, *103*, 491–504.

Thoma, J. (ms). Merely Means Paternalist? Prospect Theory and 'Debiased' Welfare Analysis. Unpublished manuscript.

Thorstad, D. (2019). Permissive Metaepistemology. *Mind*, *128*(511), 907–926.

Titelbaum, M. G., & Kopec, M. (ms). Plausible Permissivism. Unpublished manuscript.

van Fraassen, B. C. (1984). Belief and the Will. *Journal of Philosophy*, *81*, 235–56.

van Fraassen, B. C. (1990). Figures in a probability landscape. In J. M. Dunn, & A. Gupta (Eds.) *Truth or Consequences*, (pp. 345–56). Dordrecht: Kluwer.

van Fraassen, B. C. (1995). Belief and the Problem of Ulysses and the Sirens. *Philosophical Studies*, *77*(1), 7–37.

von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 2$^{\text{nd}}$ ed.

Wald, A. (1945). Statistical decision functions which minimize the maximum risk. *Annals of Mathematics*, *46*(2), 265–280.

Williams, J. R. G. (2012). Gradational Accuracy and Non-Classical Semantics. *Review of Symbolic Logic*, *5*(4), 513–537.

Williamson, J. (2010). *In Defence of Objective Bayesianism*. Oxford: Oxford University Press.

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

Zollman, K. J. S. (2010). The Epistemic Benefit of Transient Diversity. *Erkenntnis*, *72*, 17–35.