Doomsday and objective chance

Teruji Thomas

Global Priorities Institute | February 2021

GPI Working Paper No. 2-2021

Doomsday and Objective Chance

Teruji Thomas*

Abstract

Lewis's Principal Principle says that one should usually align one's credences with the known chances. In this paper I develop a version of the Principal Principle that deals well with some exceptional cases related to the distinction between metaphysical and epistemic modality. I explain how this principle gives a unified account of the Sleeping Beauty problem and chance-based principles of anthropic reasoning. In doing so, I defuse the Doomsday Argument that the end of the world is likely to be nigh.

1 Introduction

It's often the case that one should align one's credences with what one knows of the objective chances.[1] Lewis (1980) calls this the *Principal Principle*. For example, it is often the case that if one knows that a fair coin has been tossed, then one should have credence $1/2$ that heads came up. The standard caveat—the reason for the 'often'—is that one

*Global Priorities Institute, University of Oxford; teru.thomas@oxon.org. This is a draft paper dated February 2021; please check for updates before citing. I am especially grateful to David Manley for discussing various background issues with me, and to Natasha Oughton and Elliott Thornley for research assistance.

[1] This paper is mainly a project in Bayesian epistemology, and I'll speak throughout about what one 'knows' as a shorthand for what evidence one has in the sense relevant to Bayesian conditionalization. This is a natural way of speaking, but nothing turns on the identification of evidence with knowledge.

sometimes knows too much to simply defer to the chances. A trivial example: once one sees that the coin has landed tails, one should no longer have credence 1/2 in heads. In such cases, one has what Lewis calls 'inadmissible evidence'.

In this paper, I develop a version of the Principal Principle that handles two subtler kinds of exceptions, both related to the distinction between epistemic and metaphysical modality. The first arises because one can know some contingent truths *a priori*. The second is related to the fact that even an ideal thinker may be ignorant of certain necessary truths—in particular, one may not know who one is.

The second type of case is my main focus, and I will illustrate it with two well-known examples: the Sleeping Beauty puzzle (Elga, 2000) and the Doomsday Argument (Leslie, 1992). My version of the Principal Principle, labelled simply PP, yields standard views about both these cases: it yields the thirder solution to Sleeping Beauty, and denies that Doomsday is especially close at hand. These conclusions are well represented in the literature; my contribution is to present them as an attractive package deal, following from a single principal about the conceptual role of objective chance. The Doomsday Argument, in particular, is usually analysed in quite different terms, using anthropic principles like the Strong Self-Sampling Assumption and the Self-Indication Assumption. I will explain how PP leads to chance-based versions of these assumptions, unified in a principle I call Proportionality.

In §2, I introduce the existing version of the Principal Principle that will be my starting place. In §3, I explain the problem that arises from *a priori* contingencies, and suggest a preliminary solution. In §4, I explain how this preliminary solution faces the problem of self-locating ignorance. I state my preferred principle, PP, and show how it handles Sleeping Beauty and Doomsday. In §5, I state the principle of Proportionality and compare it to the standard anthropic principles. (The proof of the main result is in the appendix.) In §6, I briefly consider what my chance-based principles suggest about anthropic reasoning based simply on *a priori* likelihood, rather than chance. Section 7 sums up and points out one remaining difficulty for my theory.

Along the way, I will use the framework of epistemic two-dimensionalism (Chalmers, 2004) to model the connection between epistemic and metaphysical modality. I won't be defending two-dimensionalism in this paper, but it does conveniently represent the phenomena with which I am concerned. My hope is that its critics can find equivalent (or better!) things to say in their own frameworks.

2 Background

I will think of the Principal Principle as a constraint of rationality on an agent's *ur prior*. This ur prior, which I denote Cr, is a probability measure reflecting the agent's judgments of *a priori* probability and evidential support. Or, better, not a probability measure but a Popper function, a two-place function directly encoding conditional probabilities.[2] I'll refer to the arguments of Cr as 'hypotheses'. 'Propositions' would also do, but I use different terminology to emphasize that hypotheses are individuated hyperintensionally: the hypothesis that water is H_2O is distinct from the hypothesis that water is water, and someone could reasonably have different credences in them.

What's the relationship between ur priors and credences? Suppose that at time t one has total evidence E and credence function Cr_t. Then one should (I suggest) satisfy the norm of

Ur Prior Conditionalization. $\mathrm{Cr}_t(H) = \mathrm{Cr}_E(H) := \mathrm{Cr}(H \mid E)$.

As is well known, Ur Prior Conditionalization entails ordinary Bayesian Conditionalization: if one's evidence strengthens from E to $E \,\&\, E'$, then one's credences change from $\mathrm{Cr}_E(H)$ to $\mathrm{Cr}_E(H \mid E')$. However, Ur Prior Conditionalization has the advantage that it handles situations where one's evidence changes in other ways, like cases of forgetting: whatever happened in the past, the appropriate thing now is to conditionalize one's ur prior on one's current evidence. The question of whether Ur Prior Conditionalization handles such cases *correctly* will be

[2]See Hájek (2003) for reasons one might take conditional probabilities as primitive. Unconditional probabilities can be recovered as probabilities conditional on a tautology.

relevant later on, but for the most part I will treat this as a working hypothesis to which I do not know any comparably adequate alternative.[3]

Now, as to the Principal Principle, I will start from a version developed by Meacham (2010) and (as he notes) in unpublished work by Arntzenius. Letting $\langle \mathrm{ch}(H \mid E) = p \rangle$ stand for the hypothesis that the chance of H, given E, is p, Arntzenius's formulation of the principle is

$$\mathrm{Cr}(H \mid E \,\&\, \langle \mathrm{ch}(H \mid E) = p \rangle) = p.$$

A more general claim will also be useful. Let $\langle \mathrm{ch} = f \rangle$ stand for the hypothesis that the chances agree with the (perhaps only partially defined) Popper function f; thus $\langle \mathrm{ch} = f \rangle$ is effectively a conjuction of hypotheses of the form $\langle \mathrm{ch}(H \mid E) = p \rangle$. I will write Cr_f for the Popper function obtained by conditionalizing Cr on $\langle \mathrm{ch} = f \rangle$:

$$\mathrm{Cr}_f(H \mid E) := \mathrm{Cr}(H \mid E \,\&\, \langle \mathrm{ch} = f \rangle).$$

Then the general principle I attribute to Meacham and Arntzenius is

PP1. $\mathrm{Cr}_f(H \mid E) = f(H \mid E)$.

More precisely, the two sides should be equal *when both are defined*, but from now on I'll always leave out this type of qualification.[4]

I defer to Meacham for a careful explanation of the connection between PP1 and Lewis's classic version of the Principal Principle, but two points are especially relevant. First, PP1 is compatible with the existence of non-trivial chances even in worlds where the fundamental physics is deterministic. For example, if E is a suitable macroscopic specification of the initial conditions of a fair coin toss, and H is the hypothesis that the coin lands heads, we may well have $\mathrm{ch}(H \mid E) = 1/2$. This doesn't contradict the claim of determinism that, if E' is a complete *micro*physical specification of the initial conditions, then either $\mathrm{ch}(H \mid E') = 1$ or $\mathrm{ch}(H \mid E') = 0$. So I won't hesitate to treat coin tosses as genuinely chancy.

[3]See e.g. Moss (2015, pp. 174–176) for discussion of Ur Prior Conditionalization, and Titelbaum (2016) for some relevant alternatives.

[4]To clarify the connection to Meacham's work: the hypothesis $\langle \mathrm{ch} = f \rangle$ takes the place of what he calls a 'chance-grounding' proposition.

The second important point is that, unlike one of Lewis's formulations, PP1 does not need an exception for inadmissible evidence. Continuing the example from the previous paragraph, suppose that the agent learns that H is true. Then, for any Popper function f, PP1 gives

$$\mathrm{Cr}_f(H \mid H \,\&\, E) = f(H \mid H \,\&\, E) = 1.$$

So after one learns the result of the coin toss, one is no longer bound to give credence 1/2 to heads.

3 The Principal Principle and A Priori Contingents

The first problem for PP1 arises from the distinction between epistemic and metaphysical modality, and in particular from the phenonenon of *a priori* contingents.

> **Example 1: Topper Comes Up.** Suzy is about to flip a coin, which she knows to be fair. She introduces 'Topper' to rigidly designate whichever side of the coin will come up. Because of the way she introduces the term, she can be certain that Topper comes up. However, Topper is either heads or tails. Suzy knows that, either way, there is a 1/2 chance that Topper comes up. So her credence that Topper comes up should not equal the known chance that Topper comes up.[5]

If E is a suitable specification of the coin-tossing set-up, and *Top* is the hypothesis that Topper comes up, then

$$\mathrm{Cr}(\mathit{Top} \mid E \,\&\, \langle \mathrm{ch}(\mathit{Top} \mid E) = 1/2 \rangle) = 1$$

contradicting PP1. This example trades on the idea that chance has to do with metaphysical or nomological modality, whereas credence is a matter of epistemic modality. It's essentially *a priori* for Suzy that Topper comes up, and that's why Suzy gives it credence one. But it's not necessary that Topper comes up, and so too it's not chance one.

[5]This example is inspired by a similar one in Hawthorne and Lasonen-Aarnio (2009, pp. 95–96).

Similar problems can arise for natural kind terms. Suppose that 'water' rigidly designates what one might describe for short as the predominant wet stuff (which turns out to be H_2O). Then we can dream up a case in which it's *a priori* that the predominant wet stuff is water, and yet there's a 1/2 chance that the predominant wet stuff is H_2O. This would again enable a counterexample to PP1.

Finally, similar cases arise for indexicals.

> **Example 2: The Sheds.** I'm Carlos; Ramon is my twin. There are two windowless sheds. A fair coin is tossed. If heads, Ramon goes in Shed 1 and I go in Shed 2; if tails, the other way around. We sit there in the dark. Just before noon, partial amnesia is induced: although we both remember the general set-up, neither of us is sure whether he is Carlos or Ramon, nor how the coin landed, nor whether he is in Shed 1 or Shed 2.
>
> If I'm Carlos and this is Shed 1, then the chance that I'm in this shed is the chance that Carlos is in Shed 1, i.e. 1/2. Similarly if I'm Ramon and this is Shed 2, and so on. In any case, there's a 1/2 chance that I'm in this shed. And yet I'm certain that I am in this shed.

3.1 Neutrality

To avoid the problems raised by these examples, we could simply restrict the Principal Principle to cases in which the relevant hypotheses do not involve proper names, or natural kind terms, or indexicals, or anything of the sort—in short, to the kind of hypotheses that Chalmers (2011) calls *neutral*:[6]

PP2. If E and H are neutral hypotheses, then $\mathrm{Cr}_f(H \mid E) = f(H \mid E)$.

[6] Because of the conditionalization, it really suffices that E and $H \,\&\, E$ are neutral. While I won't focus on this issue here, Lewis (1980, pp. 268–9) essentially points out that the Popper function f must also be given in a suitably neutral form. If I know that the chance of heads is x, and, unknown to me, x equals 1/4, then I'm under no compulsion to set my credence in heads to 1/4.

While this basic proposal will require some amendment in §4, its meaning and limitations will be clearer if we pause, first to explain how the neutrality restriction works within the framework of *epistemic two-dimensionalism*, (Chalmers, 2004, 2011), arguably its natural home; and second to explain how PP2 purports to give a full account of Topper Comes Up and similar cases.

Recall that the *intension* of a hypothesis is a set of possible worlds—the worlds that would make the hypothesis true. Two hypotheses are necessarily equivalent iff they have they same intension. My assumption that chance is a form of metaphysical modality amounts to the claim that the chance of a hypothesis depends only on its intension. However, rational credences can distinguish between necessarily equivalent hypotheses. For example, suppose that Suzy's coin in fact lands heads-up. Then the hypothesis that Topper comes up is necessarily equivalent to the hypothesis that heads comes up. Yet Suzy gives them different credences.

To represent the distinctions made by rational credences, two-dimensionalists introduce a second dimension of *epistemic scenarios*. These are like possible worlds, but individuated by epistemic criteria. For example, there are some scenarios in which Topper is heads, and others in which Topper is tails; as Suzy's uncertainty attests, these are distinct and genuine epistemic possibilities. The *primary intension* of a hypothesis is the set of scenarios (rather than possible worlds) in which it is true; two hypotheses are *a priori* (rather than necessarily) equivalent iff they have the same primary intension.

For my purposes, the key point is that, according to Chalmers, each scenario picks out (i) a possible world as actual; and (ii) an intension, i.e. a set of possible worlds, for each hypothesis. For example, some scenarios pick out a world in which heads comes up. In such a scenario, Topper is heads, and the intension of *Top* is the set of worlds in which heads comes up. Other scenarios pick out a world in which tails comes up. Then Topper is tails, and the intension of *Top* is the set of worlds in which tails comes up.[7]

[7] If heads actually comes up, there is no possible world in which Topper is tails. However, there are possible worlds in which tails comes up, and the thought is that,

Chalmers calls a hypothesis *neutral* if and only if, unlike *Top*, it has the same intension in every scenario.

We can now see why the restriction to neutral hypotheses, understood in this way, avoids the problems raised by Topper Comes Up and similar cases. Ur priors can't distinguish between *a priori* equivalent hypotheses, like 'Topper comes up' and 'Whichever side comes up, comes up'. PP1 is bound to fail insofar as such hypotheses are not necessarily equivalent, so that the chances *can* distinguish them. This problem cannot arise for the class of neutral hypotheses, however: one can show that two neutral hypotheses that are *a priori* equivalent are also *a priori* necessarily equivalent (they have the same intension in each and every scenario), and therefore *a priori* have the same chance.[8]

Even if the restricted principle PP2 avoids problems, one might worry that it applies too rarely to constrain credences in all the expected ways. In one respect, which I'll discuss in §4, this turns out to be a very serious worry, but I think it is worth having a first-pass explanation of how PP2 could give the desired results in a case like Topper Comes Up. Let us focus on Suzy's credence that Topper is heads. Since this hypothesis is not neutral, PP2 does not directly tell us the right credence. However, we can reason in two stages. First, the hypothesis that heads comes up is more plausibly neutral, and, if so, PP2 does require Suzy's credence in heads to be 1/2. Second, it's *a priori* for Suzy that Topper is heads iff heads comes up. Therefore Suzy must also have credence 1/2 that Topper is heads.

Not only do we get the right conclusion, the explanation for it strikes me as perspicacious. At any rate, it illustrates that the restriction to neutral hypotheses is not debilitating insofar as there are what I'll call *neutral paraphrases* of more general hypotheses. Here, E° is a neutral paraphrase of E if and only if E° is neutral and E and E° are *a priori* equivalent. Because of this last condition, E and E° are interchangeable when it comes to ideal ur priors. For example, 'heads comes up' is a neutral paraphrase of 'Topper is heads'. Suzy's credence

in any *scenario* that picks out such a world as actual, Topper is tails.

[8] This depends on a seemingly harmless assumption, adopted by Chalmers, that every possible world is actual in some scenario.

in the latter is determined by her credence in the former, which is in turn determined by PP2.

4 The Principal Principle and Self-Location

4.1 The Problem

While it is arguable that a wide range of hypotheses about the world do admit neutral paraphrases, it is unfortunately impossible to maintain that our evidence in ordinary circumstances—circumstances in which we expect the Principal Principle to be binding—is of that type. The reason is that neutral hypotheses exclude the use of indexicals. In a scenario where I am Carlos, the intension of the hypothesis *I am sitting* consists of the worlds in which Carlos is sitting; in a scenario where I am Ramon, it picks out the worlds in which Ramon is sitting. Thus a neutral hypothesis can arguably give an adequate third-personal, qualitative description of the world, but it can do nothing to identify one's own situation. Suppose, for example, that in Topper Comes Up, Suzy knows it's noon. Even though this is a perfectly ordinary thing to know, PP2 will not entail that Suzy's credence in heads should be $1/2$, because her evidence cannot be given a neutral paraphrase. In fact, PP2 only applies if one has essentially no knowledge whatsoever of what one is like or where one is in space and time.

We can use the two-dimensionalist framework to shed some light on this situation. Following Chalmers again, we can identify each epistemic scenario with a *centered possible world*: a triple (w, x, t) where w is a possible world, and x is an individual and t a time in w. I'll refer to (w, x, t) as a *centering* of w, and (x, t) as a *center*. Thus the primary intension of a hypothesis is a set of centered worlds. For example, the primary intension of the hypothesis *I'm sitting in a comfy chair* consists of the centered worlds (w, x, t) such that, *a priori*, if I'm x at t in w, then I'm sitting in a comfy chair.[9] Now, it may be that some formally

[9]The identification of scenarios with centered worlds, and the question of whether this is fully appropriate, are somewhat delicate; I defer to Chalmers (2011) for discussion. The use of centred possible worlds to model self-locating ignorance is standard since at least Lewis (1979), and most of the rest of this paper could be written in a

possible centered worlds do not represent genuine epistemic possibilities, even *a priori*. Perhaps it is *a priori* for me that I am not a rock; then (w, x, t) does not correspond to an epistemic scenario, if x is a rock at time t in w. When I talk about centered worlds, I only mean those that represent genuine epistemic possibilities.

Now back to the immediate point. Let w be some possible world. For a hypothesis to be neutral, it must have the same intension with respect to every scenario, and in particular with respect to every centering of w. It follows that if the primary intension contains one centering of w, it must contain them all. This makes precise the idea that neutral hypotheses (or those with neutral paraphrases) completely fail to locate the subject in the world. In contrast, the primary intension of the hypothesis that it's noon contains only centered worlds (w, x, t) such that it's noon where x is at time t. Such ordinary evidence cannot be given a neutral paraphrase.

4.2 Self-Locating Hypotheses

One strategy would be to supplement PP2 by principles of a different kind that together constrain Suzy's credences in the right way. I will consider some such principles in §5. However, this strategy seems back-to-front. The Principal Principle, whatever the details, is supposed to express a platitude about how knowledge of the chances ordinarily constrains our credences. It hardly matters what it says about bizarre cases of complete indexical ignorance; it ought to apply directly in situations that at worst idealize what we take to be the ordinary case.

I propose instead to formulate a modification of PP2 that applies directly when one *does* have fully self-locating evidence: that is, more carefully, when the primary intension of one's evidence contains at most one center for each possible world.

Lest this appear a radical move, let me emphasize that it is a natural interpretion of what Lewis (1980) himself says. He develops the Prin-

---

Lewisian framework. Note though that Lewis claims the objects of belief are properties, whose intensions are sets of centred worlds. In contrast, for two-dimensionalists, the (ordinary, not primary) intension of a hypothesis is still a set of possible worlds. See Magidor (2015) for critique especially of the Lewisian tradition.

cipal Principle in a setting where one's credences assign probabilities to possible worlds, and therefore do not explicitly distinguish different centers within each world. However, the point is *not* that his principle applies only in bizarre cases of complete self-locating ignorance! He applies it to ordinary coin-tossing cases, after all. Rather, uncentered possible worlds usually suffice because each such world comes with an *implicit* center, picked out by the agent's self-locating evidence. As Lewis says, we only need to use centers explicitly if we want to handle cases in which 'one's credence might be divided between different possibilities within a single world' (Lewis, 1980, p. 268). So Lewis's principle applies when one's credences are *not* so divided, i.e. when one has fully self-locating evidence. Moreover, it applies no matter what the implicit centerings may be. My aim is to spell out this picture in detail, as Lewis does not.

To emphasize the role of indexicals, I will often represent potentially non-neutral hypotheses in the form $\langle \text{I am } G \rangle$. $\langle \text{I am } FG \rangle$ means $\langle \text{I am } F \rangle$ & $\langle \text{I am } G \rangle$, and so on. I'll say that a hypothesis is *fully self-locating* iff its primary intension contains at most one centering of each possible world. The proposal is to restrict the Principal Principle to cases of fully self-locating evidence. However, there is a more convenient way to put this. Say that $\langle \text{I am } G \rangle$ is *merely self-locating* relative to background evidence E if it picks out exactly one centering of each world compatible with E. More carefully, I am talking about primary intensions, so the condition is that, if the primary intension of E contains a centering of w, then the primary intension of E & $\langle \text{I am } G \rangle$ contains *exactly one* centering of w. It follows that E & $\langle \text{I am } G \rangle$ is fully self-locating.

In these terms, the main proposal of this paper is that the chances bind credences conditional on each merely self-locating hypothesis:

PP. If E and H are neutral hypotheses, and $\langle \text{I am } G \rangle$ is merely self-locating relative to E, then

$$\mathrm{Cr}_f(H \mid E \,\&\, \langle \text{I am } G \rangle) = f(H \mid E).$$

The restriction to neutral E and H is still important here, but in §5

I will develop a less restricted principle—Proportionality—as a consequence of PP.

One might worry that ordinary evidence is never *fully* self-locating: perhaps it does not narrow things down to exactly one individual and one time in each world compatible with one's evidence. I'll consider a troubling form of this worry in §7, but for now I will just address the most mundane form: one's evidence may not often pin down a precise time. There are two basic responses.

The first is that one can have fully self-locating evidence even if one does not know what time it is in the ordinary sense that one does not know what clocks are saying right now. Clockfaces are only one way of picking out an instant in each world.

However, one may still worry that one's evidence is coarse-grained in a way that just can't pin the present down exactly. There may be some deep issues here about perception and even about the metaphysics of time, but the short answer is that we are allowed, as far as PP goes, to count times in a coarse-grained way. We need not take 'one time' to mean 'one instant' rather than 'one interval of unit length', where the units are adjustable and we count non-overlapping unit-length intervals as different times. What's crucial in applying PP is that the precision with which $\langle$I am $G\rangle$ locates me in the world is independent of how the world turns out, conditional on E.

4.3 Sleeping Beauty

To see PP in action, consider this famous example:[10]

> **Example 2: Sleeping Beauty.** On Sunday night, Beauty knows she is in the following situation. After she goes to sleep, a fair coin will be tossed. She will be awakened on Monday. A few minutes later, she will learn it is Monday. Then she will go back to sleep. If the coin landed heads, she will sleep through Tuesday. But if it landed tails, her memories of Monday will be erased, and she will

[10]The example was made popular by Elga (2000); see his first footnote for its history.

> be awakened on Tuesday morning. Thus, when Beauty wakes on Monday, she does not know whether the coin landed heads or tails, nor, supposing the coin landed tails, whether it is Monday or Tuesday.
>
> What should Beauty's credence in heads be (a) on Sunday night; (b) on first waking; (c) after learning it is Monday?

PP allows us to analyse the case as follows. On Sunday night, Beauty's evidence, as normal, is fully self-locating. Therefore PP applies and tells us she should have credence 1/2 in heads. So too *after* Beauty learns it's Monday. On first waking, however, her evidence is not fully self-locating, and PP does not apply. Nevertheless, we can argue from PP that she must have credence 1/3 in heads. Consider the three hypotheses *HM* (heads, which implies it's Monday), *TM* (tails and it's Monday), and *TT* (tails and it's Tuesday). Assuming that Beauty will update by conditionalization on learning it's Monday (i.e. $HM \vee TM$), she must already give *HM* and *TM* the same credence. Now consider what would happen were she instead to learn $HM \vee TT$. She would again have fully self-locating evidence, and should again have credence 1/2 in heads. So she must already give *HM* and *TT* the same credence. All together, she gives the same credence to each of the three hypotheses *HM*, *TM*, and *TT*. Since these are mutually exclusive and exhaust the possibilities open to her, she must give credence 1/3 to each.

This pattern of credences is called the 'thirder' position in the literature on Sleeping Beauty. I find the extant arguments for thirderism quite compelling, and I am happy to refer to them as corroboration for my view. However, the analysis I've presented is slightly different from the most common way of understanding thirderism. Elga (2000) appeals to a principle of indifference: Beauty should, on waking, consider the hypotheses *TM* and *TT* equally likely, since her evidence is fully symmetric between them. But this suggestion invites standard worries about indifference reasoning, including the thought that Beauty might have symmetrical but only imprecise credences in these hypotheses (Weatherson, 2005). My argument is different, and isn't directly susceptible to such worries. Instead of appealing to evidential symmetry, I claim that Beauty should align her credences with the known chances, not only after she learns it's Monday, but also if she were instead to learn $HM \vee TT$, on the basis that these are both merely self-locating hypotheses relative to her other evidence.

Of course, thirderism is not the only standard position when it comes to Sleeping Beauty. As Elga explains, the main rivals to thirders are *halfers*, who claim that Beauty should give credence 1/2 to heads when she wakes up, as well as on Sunday night. I can't do justice to the whole literature, and want to focus on my positive proposal, but it seems significantly more difficult to do for halferism what I've done for thirderism here: to embed it in a package that includes systematic norms for updating (as in Ur Prior Conditionalization) and a natural version of the Principal Principle (as in PP). Beauty's evidence after learning it's Monday is structurally very similar to her evidence on Sunday night, so it's hard to see why the Principal Principle would apply in the second case but not the first. On the other hand, if, as 'double halfers' claim, Beauty should have credence 1/2 in heads at all three times, then she must not apply Bayesian conditionalization when she learns it's Monday.[11]

4.4 The Doomsday Argument

Here is another example. It is very similar to Sleeping Beauty, but it will be useful to consider it separately, because it is commonly analysed using quite different tools, which I will contrast with PP in section 5.

> **Example 3: Doomsday.** There's a 1/2 chance that humanity goes extinct at an early stage, resulting in a total of 100 billion human beings who ever live (call this outcome *early doom*); and a 1/2 chance that humanity hangs on much longer, resulting in 100 quadrillion human beings who ever live (call this outcome *late doom*). I'm human. Against this evidential background, I learn that I am the

[11] On the first point, Lewis (2001) claims that Beauty has inadmissible evidence once she learns it's Monday, but it seems hard to independently justify this claim. On the second, see Titelbaum (2016) for a survey of alternative updating methods and their problems.

70 billionth human to be born. What should my credence be in early doom?

As the *Doomsday Argument* notes, knowing I am the 70 billionth human rules out many possibilities that are compatible with both early doom and late doom, but vastly many more that are only compatible with late doom (for example, the possibility that I am the 200 billionth human). So, for any reasonable priors, that piece of evidence should result in a dramatic shift in credence towards early doom. Unless I was antecedently ridiculously confident in late doom—*far* more confident than the stated 1/2 chance—I should now be almost certain of early doom.[12]

The example is practically significant because our actual evidential situation is stylistically similar to the one described. We have some idea about the various kinds of extinction risks we face (either as a species or as a global ecosystem), and a fairly precise idea of how far along we are since life began. The basic logic of the Doomsday Argument generalizes to more complicated cases, and *seems* to show that an early doom for humanity is much more likely (epistemically speaking) than the chances would on their face suggest.

However, in parallel to my analysis of Sleeping Beauty, PP implies that my posterior credence in early doom should be 1/2. At least, it does so for reasonable ways of filling in the details. Most simply, assume that everyone has the same lifespan; then the hypothesis that I'm the 70 billionth human is fully self-locating by the criteria sketched in §4.2. Thus it is after, not before, learning that I am the 70 billionth human that PP binds my credences to the chances. This, along with Ur Prior Conditionalization, commits me to having been 'ridiculously' confident in late doom prior to gaining the new evidence. But, then again, prior to that evidence I was in the ridiculous epistemic state of having essentially

[12]This is a simple version of the Doomsday Argument treated explicitly by Leslie (1992) and attributed to Brandon Carter. See Bostrom (2002) for a discussion of its history. Note that your current evidence may well be fully self-locating even if you have little idea of your birth-rank among humans (cf. my discussion of knowing the time in §4.2). So this Doomsday Argument says nothing about what should happen if you were to learn your birth-rank in real life.

no self-locating information. We shouldn't be too worried about getting surprising results about such exotic epistemic positions.

5 The Principal Principle and Anthropic Reasoning

5.1 Proportionality

By design, PP only directly constrains the credences of agents with fully self-locating evidence. But, as already hinted in my analysis of Sleeping Beauty and Doomsday, it has broader implications. I'll now draw out some of those implications, and show how they improve upon the anthropic principles that are commonly used to analyse Doomsday.

Here is the main result. Consider hypotheses E and $\langle \text{I am } G \rangle$. Very roughly, I will use $N_f(G \mid E)$ to denote the expected number of things that are G, conditional on E. More carefully, recall that the primary intension of $\langle \text{I am } G \rangle$ contains zero or more centerings of each possible world. Then I define $N_f(G \mid E)$ to be the expected number of such centerings, according to $f(- \mid E)$. So, for each world w, we take the number of centerings of w in the primary intension of $\langle \text{I am } G \rangle$, we multiply that by the probability (according to f, conditional on E) that w is actual, and then we sum over worlds.[13] Thus $N_f(G \mid E) = 1$ if $\langle \text{I am } G \rangle$ is merely self-locating with respect to E, and will be higher insofar as $\langle \text{I am } G \rangle$ fails to pin down my location.

I claim that PP entails the following principle, given a sufficiently rich domain of hypotheses; the proof is in the appendix.

Proportionality. Suppose E is a neutral hypothesis. Then

$$\mathrm{Cr}_f(\langle \text{I am } F \rangle \mid \langle \text{I am } G \rangle \,\&\, E) = \frac{N_f(FG \mid E)}{N_f(G \mid E)}.$$

Note that (unlike in PP2) the restriction to neutral E is not onerous, since the overall evidence $\langle \text{I am } G \rangle \,\&\, E$ is effectively arbitrary. Proportionality is a sophisticated version of the intuitive idea that my credence

[13] This recipe is a little rough for the usual reason that there may be uncountably many relevant worlds, and we can't just sum over them; I'll give a more formal definition in the appendix.

that I'm F, given that I'm G, should be high insofar as most Gs are Fs.[14] I'll draw out its precise meaning by comparing Proportionality to two somewhat similar principles that are standard in the literature.

5.2 The Self-Sampling Assumption

The first of the two main anthropic principles is, in Bostrom's influential formulation, the

> **Strong Self-Sampling Assumption (SSSA).** One should reason as if one's present observer-moment were a random sample from the set of all observer-moments in its reference class.[15]

Here an 'observer-moment' is what I have been calling a centered possible world: one's *present* observer moment is the actual world centered on oneself and the present time. Although SSSA is not precisely stated, the gist is that one should consider different merely self-locating hypotheses to be equally likely. So, for example, Beauty should be indifferent between Monday and Tuesday, conditional on tails. In Doomsday, the idea is that I should initially give equal credence to different hypotheses about my birth-rank in each world separately. Because of this, my initial credence that I'm the 70 billionth human is a million times higher conditional on early doom than on late doom. This determines how strongly I should update in favour of early doom upon learning my birth-rank: the subjective odds of early doom increase by a factor of one million.

My own analysis of Doomsday used PP to determine my posterior credence in early doom directly. It is unnecessary to adduce SSSA as a separate principle, since the following version of it is a simple application of Proportionality:

[14] Proportionality is closely related to what Manley (2014) calls 'Typicality', but importantly different from what Arntzenius and Dorr (2017) call 'Proportion': roughly, the latter requires the stated credence to equal the expected proportion of Gs that are Fs.

[15] Bostrom (2002, p. 162). The (not 'Strong') Self-Sampling Assumption applies to observers, rather than observer moments, but that won't help with Sleeping Beauty cases, and is actually incompatible with SSSA.

Uniformity. If $\langle \text{I am } G \rangle$ and $\langle \text{I am } G' \rangle$ are merely self-locating relative to a neutral hypothesis E, then

$$\mathrm{Cr}_f(\langle \text{I am } G \rangle \mid E \,\&\, \langle \text{I am } G \text{ or } G' \rangle) \\ = \mathrm{Cr}_f(\langle \text{I am } G' \rangle \mid E \,\&\, \langle \text{I am } G \text{ or } G' \rangle).$$

Besides being much more precise, Uniformity differs from SSSA in several important respects.

First, Uniformity only applies conditional on an appropriate chance hypothesis $\langle \mathrm{ch} = f \rangle$. I'll say more about this limitation in §6.

Second, Uniformity makes sense even when some worlds compatible with E include infinitely many observer-moments, whereas there is no entirely reasonable way to randomly sample from an infinite set.[16]

Third, as usually conceived, SSSA is a principle of indifference between different merely self-locating hypotheses, similar to the indifference principle Elga used to analyse Sleeping Beauty. In contrast, Uniformity is based on a claim about the applicability of the chance-credence link. Of course, PP does include a kind of indifference claim, to the effect that *all* merely self-locating hypotheses are equally good from the point of view of the Principal Principle.

A fourth, closely related difference is that SSSA appeals to the the idea of a 'reference class' of observer-moments. Uniformity treats all merely self-locating hypotheses as equally good, without limitation to a narrower reference class (but with the understanding that centered worlds include only genuine *a priori* possibilities). Bostrom uses flexibility in the choice of reference class to resolve various problems that arise from his theory, including the Doomsday Argument. This flexibility seems unnecessary when it comes to Uniformity: PP treats the Doomsday Argument without further recourse to reference classes.

5.3 The Self-Indication Assumption

The second, more controversial anthropic principle is the

[16] I don't claim to solve all the related problems that arise from infinite worlds, for discussion of which see Bartha and Hitchcock (1999b), Weatherson (2005), and especially Arntzenius and Dorr (2017). It's worth mentioning that Popper functions need not be countably additive.

> **Self-Indication Assumption (SIA).** Given the fact that you exist, you should (other things equal) favor hypotheses according to which many observers exist over hypotheses on which few observers exist. (Bostrom, 2002, p. 66)

This is again rather imprecise, but SIA is commonly understood as a claim about the evidential import of the fact that one exists: conditionalising one's ur prior on that evidence increases the relative likelihood of worlds with large populations. This idea is especially clearly stated by Bartha and Hitchcock (1999a), but goes back to Dieks (1992).

One *post hoc* motivation for SIA is that it provides a way of blocking the Doomsday Argument. Suppose that we think the chance-credence link is properly given by PP2. Then, knowing only the chance hypothesis stated in Doomsday, I should have a 1/2 credence in early doom and 1/2 in late doom. For the sake of discussion, let's also suppose that, compatible with these credences, I know that I'm human if I exist at all. Next, I conditionalize on two pieces of evidence: (E_1) that I exist, and (E_2) that I am, specifically, the 70 billionth human. The Doomsday Argument really shows us that *given* E_1, E_2 shifts my credences dramatically towards early doom. But SIA tells us that conditionalizing on E_1 itself shifts my credences towards *late* doom. So if we interpret SIA in exactly the right way, these two shifts will cancel out, and the net effect of learning E_1 and E_2 is to leave my credence in early doom at the original 1/2.

Is there any independent reason to think that E_1 has exactly the evidential significance required? Bartha and Hitchcock (1999a, p. 349) provide what they call a 'just-so story': if the 100 billion people in the early doom world and the 100 quadrillion people in late doom world were chosen separately and uniformly at random from a stock of possible people, then any one of those possible people would have a greater chance (and greater to just the right degree!) of being selected into the late doom world. But even if we managed to take this just-so story seriously as a piece of cosmology, the upshot would be unclear. How does it help with cases of self-location within a life, as in Sleeping Beauty? And notice that the metaphysical claim that the population is chosen at random is compatible with the not unreasonable epistemic claim that it is *a priori*, for me, that I exist. But if it is *a priori*, then it has no evidential weight for me at all, contrary to SIA. Even the just-so story equivocates between metaphysical and epistemic modality in the way I have been trying to avoid.

Nevertheless, there is a precise sense in which Proportionality requires one to give more credence to large-population hypotheses than the chances naively suggest. It entails:

Weighting. If E and H are neutral hypotheses, then

$$\mathrm{Cr}_f(H \mid E \,\&\, \langle \text{I am } G\rangle) = \frac{N_f(G \mid H \,\&\, E)}{N_f(G \mid E)} f(H \mid E).$$

Weighting is a precise generalization of the claims that, before learning it's Monday, Beauty should be quite confident in tails, and that, before learning I'm the 70 billionth human, I should be extremely confident in late doom.

6 Beyond Chance

This paper has been about objective chance, and the anthropic principles developed in §5 are formulated in terms of a chance hypothesis $\langle \mathrm{ch} = f\rangle$. As I mentioned in §3, my understanding of chance-talk is pretty broad: it's not just limited to indeterministic interpretations of quantum mechanics, or anything like that. Still, I agree that there are situations where talk of chances would seem misplaced, including cases in which we are considering the relative plausibility of different scientific theories. So I don't claim to have recovered the full scope of the anthropic principles that have been proposed in the literature. But I have shown that one can get pretty far with chances, and the results are suggestive of a more general analysis.

How so? Starting from an ur prior Cr, we can construct a partially defined Popper function Cr^0 that encodes judgements of evidential support *given* a background of merely self-locating evidence. Restricting ourselves to neutral hypotheses H and E, the idea is that $\mathrm{Cr}^0(H|E) = p$ holds if and only if $\mathrm{Cr}(H \mid E \,\&\, \langle \text{I am } G\rangle) = p$ whenever $\langle \text{I am } G\rangle$ is merely self-locating relative to E. With this definition, we can reformulate PP more simply as the claim that

$$\mathrm{Cr}^0(H \mid E \,\&\, \langle \mathrm{ch} = f \rangle) = f(H \mid E).$$

And the argument for Proportionality given in the appendix supports the more general claim

$$\mathrm{Cr}(\langle \text{I am } F \rangle \mid \langle \text{I am } G \rangle \,\&\, E) = \frac{N_{\mathrm{Cr}^0}(F \,\&\, G \mid E)}{N_{\mathrm{Cr}^0}(G \mid E)}$$

This generalization of Proportionality does not involve any chance hypothesis; it instead involves the judgments of evidential support represented by the Popper function Cr^0.

The point of this innovation is that sometimes our judgments of *a priori* evidential support plausibly relate to Cr^0 rather than to Cr. We just don't usually consider the case of complete self-ignorance; we take self-location for granted, as does most of the literature in epistemology that is not specifically concerned with Sleeping Beauty or Doomsday-like cases. So the loose thought that H and $\neg H$ are equally likely conditional on E may well suggest that $\mathrm{Cr}^0(H \mid E) = 1/2$ rather than $\mathrm{Cr}(H \mid E) = 1/2$. Note that $\mathrm{Cr}^0(H \mid E) = 1/2$ is what we'd expect from PP if one knew *a priori* that $\mathrm{ch}(H \mid E) = 1/2$. In that sense, the judgements reflected by Cr^0 are calibrated to the chances.

Some other ways of measuring *a priori* likelihood are at least compatible with chance-calibration. For example, one might attempt to gauge the relative likelihood of H and $\neg H$ by imagining what an angel in heaven would find plausible without having looked out to see how the universe is going.[17] But of course the angel knows perfectly well where he is, so judgments arrived at in this way must already take self-locating evidence into account.

For illustration, consider a version of Doomsday in which early doom and late doom are supposed to be equally likely *a priori*, but this isn't cashed out in terms of chances. If 'equal likelihood' is understood in terms of Cr, then (setting aside SIA and other shenanigans)

[17] See Bostrom (2002, pp. 32ff) for a similar heuristic.

the Doomsday Argument does seem to show that someone with fully self-locating evidence will be dramatically more confident in early doom than in late. But this point is not very interesting unless we have a decent grip on what is epistemically likely given the ridiculous evidential background of complete self-ignorance. In contrast, if 'equal likelihood' is understood in a chance-calibrated sense, or more generally just against an implicit evidential background that already includes self-locating information, then the Doomsday Argument does not go through.

7 A Final Problem

I've shown how to formulate a version of the Principal Principle that is better insulated against the problem of *a priori* contingencies and which works even in the context of self-locating ignorance. The main ideas are that one should to stick to neutral hypotheses, and that chances bind credences relative to fully self-locating evidence. The resulting picture, including Ur Prior Conditionalization, fits cleanly with the thirder view of Sleeping Beauty. It also yields chance-based versions of some well-known anthropic principles (Uniformity, Weighting, and most fundamentally Proportionality) while blocking the chance-based Doomsday Argument. Finally, one can generalize these principles beyond chances to chance-calibrated judgments of *a priori* likelihood.

The aspect of this picture that I ultimately find least satisfying is that, when it comes down to it, our ordinary evidence may *not* be fully self-locating. Given the immense size of the universe, we should take seriously the possibility that there are qualitative duplicates, or near enough, of ourselves and our surroundings somewhere else. (More carefully, the issue is that my total evidence includes in its primary intension some epistemic scenarios centered on sufficiently close duplicates of myself.) As a stylized case, consider a version of Doomsday in which the 100 quadrillion humans in the late doom world consist of a million distantly separated groups of duplicates of the 100 billion humans who would exist given early doom. Against that background, it would be hard for me to get fully self-locating evidence; reasonable evidence could at best narrow down one's identity to a million qualitatively identical people, conditional on late doom. By Weighting, I should then be extremely confident in late doom. And, to emphasise, I need not be unusually uninformed: I could be well acquainted with my environment as far as telescopes can see.

I think I have to bite the bullet here: compared to the chances, my credences should favour worlds that contain many clones of myself and my environment.[18] The consolation is that this won't interfere with ordinary applications of the Principal Principle. For example, when it comes to a fair coin toss, one should still give heads credence 1/2, so long as the expected number of one's clones doesn't depend on the toss. It is true that Proportionality, rather than PP, is more directly applicable. So once we take into account the possibility of clones, Proportionality may be the best way to think about the chance-credence link.

Appendix: Derivation of Proportionality

The argument will assume that there is a sufficiently rich space of hypotheses. Instead of formulating general conditions, here is exactly what I'll use in terms of the hypotheses E, $\langle \text{I am } F \rangle$, and $\langle \text{I am } G \rangle$.

(a) E is *non-atomic*: there is a neutral hypothesis A such that $f(A \mid E) \neq 0, 1$. It will be convenient to write A' for $\neg A$.

(b) For all integers $k \geq j \geq 0$, there is a neutral hypothesis E_{jk} whose intension contains a world w iff the primary intension of $\langle \text{I am } G \rangle \,\&\, E$ contains k centerings of w and the primary intension of $\langle \text{I am } FG \rangle \,\&\, E$ contains j centerings of w. This that allows me to formally define

$$N_f(G \mid E) = \sum_{k \geq j \geq 0} k f(E_{jk} | E) \quad N_f(FG \mid E) = \sum_{k \geq j \geq 0} j f(E_{jk} | E).$$

(c) Each E_{jk} & $\langle \text{I am } G \rangle$ has a partition by k fully self-locating hypotheses $H^1_{jk}, \ldots, H^k_{jk}$, such that $H^1_{jk}, \ldots, H^j_{jk}$ form a partition of $E_{jk} \,\&\, \langle \text{I am } FG \rangle$. It follows that each H^i_{jk} is merely self-locating with respect to E_{jk}.

[18] See Elga (2004); Weatherson (2005) for a discussion of related problems.

To proceed, choose two triples i, j, k and i', j', k'. We can apply PP:

$$\begin{aligned}\mathrm{Cr}_f(AH^i_{jk} \mid AH^i_{jk} \vee A'H^{i'}_{j'k'}) \\ &= \mathrm{Cr}_f(AE_{jk} \mid (AH^i_{jk} \vee A'H^{i'}_{j'k'}) \,\&\, (AE_{jk} \vee A'E_{j'k'})) \\ &= f(AE_{jk} \mid AE_{jk} \vee A'E_{j'k'}).\end{aligned}$$

Multiply the left and right sides by

$$\mathrm{Cr}_f(AH^i_{jk} \vee A'H^{i'}_{j'k'} \mid H) \times f(A'E_{j'k'} \mid E) \times f(AE_{jk} \vee A'E_{j'k'} \mid E)$$

where $H = E \,\&\, \langle \text{I am } G \rangle$. To simpify the result, use the identities

$$\mathrm{Cr}_f(AH^i_{jk} | AH^i_{jk} \vee A'H^{i'}_{j'k'}) \times \mathrm{Cr}_f(AH^i_{jk} \vee A'H^{i'}_{j'k'} | H) = \mathrm{Cr}_f(AH^i_{jk} | H)$$

and

$$f(AE_{jk} \mid AE_{jk} \vee A'E_{j'k'}) \times f(AE_{jk} \vee A'E_{j'k'} \mid E) = f(AE_{jk} \mid E).$$

The result is

$$\begin{aligned}\mathrm{Cr}_f(AH^i_{jk} \mid H) \times f(A'E_{j'k'} \mid E) \times f(AE_{jk} \vee A'E_{j'k'} \mid E) \\ &= \mathrm{Cr}_f(AH^i_{jk} \vee A'H^{i'}_{j'k'} \mid H) \times f(A'E_{j'k'} \mid E) \times f(AE_{jk} \mid E).\end{aligned}$$

Note that the right-hand side remains the same if we simultaneously exchange A, i, j, and k with A', i', j', and k', respectively. This must also be true of the left-hand side; therefore

$$\mathrm{Cr}_f(AH^i_{jk} \mid H) \times f(A'E_{j'k'} \mid E) = \mathrm{Cr}_f(A'H^{i'}_{j'k'} \mid H) \times f(AE_{jk} \mid E). \quad (1)$$

Here, a factor $f(AE_{jk} \vee A'E_{j'k'} \mid E)$ has been cancelled from both sides; if this factor is zero, then $f(AE_{jk} \mid E) = 0 = f(A'E_{j'k'} \mid E)$, so the equation still holds with both sides equal to zero.

If, as is always possible, we select i', j', k' so that $f(A'E_{j'k'} \mid E) \neq 0$, then we can rearrange (1) into the form

$$\mathrm{Cr}_f(AH^i_{jk} \mid H) = \alpha f(AE_{jk} \mid E) \quad (2)$$

where α is independent of i, j, k. If, instead, we select i, j, k so that $f(AE_{jk} \mid E) \neq 0$, then we rearrange (1) into the form

$$\mathrm{Cr}_f(A' H^{i'}_{j'k'} \mid H) = \beta f(A' E_{j'k'} \mid E) \tag{3}$$

where β is independent of i', j', k'. Plugging (2) and (3) into (1) shows that $\alpha = \beta$. For arbitrary $i = i'$, $j = j'$, and $k = k'$, adding (2) to (3) yields

$$\mathrm{Cr}_f(H^i_{jk} \mid H) = \alpha f(E_{jk} \mid E).$$

To determine α, recall that the H^i_{jk} form a partition of H, so that

$$\begin{aligned} 1 = \mathrm{Cr}_f(H \mid H) &= \sum_{i,j \le k} \mathrm{Cr}(H^i_{jk} \mid H) = \sum_{i,j \le k} \alpha f(E_{jk} \mid E) \\ &= \alpha \sum_{j \le k} k f(E_{jk} \mid E) = \alpha N_f(G \mid E). \end{aligned}$$

Therefore $\alpha = 1/N_f(G \mid E)$. Finally,

$$\begin{aligned} \mathrm{Cr}_f(\langle \text{I am } FG \rangle \mid H) &= \sum_{i \le j \le k} \mathrm{Cr}(H^i_{jk} \mid H) = \sum_{i \le j \le k} \alpha f(E_{jk} \mid E) \\ &= \alpha \sum_{j \le k} j f(E_{jk} \mid E) = \frac{N_f(FG \mid E)}{N_f(G \mid E)}. \end{aligned}$$

This is a restatement of Proportionality.

References

Arntzenius, F. and C. Dorr (2017). Self-locating priors and cosmological measures. In K. Chamcham, J. Barrow, S. Saunders, and J. Silk (Eds.), *The Philosophy of Cosmology*, pp. 396–428. Cambridge: Cambridge University Press.

Bartha, P. and C. Hitchcock (1999a). No one knows the date or the hour: An unorthodox application of Rev. Bayes's theorem. *Philosophy of Science 66*(3), 353.

Bartha, P. and C. Hitchcock (1999b). The shooting-room paradox and conditionalizing on measurably challenged sets. *Synthese 118*(3), 403–437.

Bostrom, N. (2002). *Anthropic Bias: Observation Selection Effects in Science and Philosophy*. Routledge.

Chalmers, D. J. (2004). Epistemic two-dimensional semantics. *Philosophical Studies 118*(1-2), 153–226.

Chalmers, D. J. (2011). The nature of epistemic space. In *Epistemic Modality*, pp. 60–107. Oxford University Press.

Dieks, D. (1992). Doomsday – or: the dangers of statistics. *The Philosophical Quarterly 42*(166), 78–84.

Elga, A. (2000). Self-locating belief and the Sleeping Beauty problem. *Analysis 60*(2), 143–147.

Elga, A. (2004). Defeating Dr. Evil with self-locating belief. *Philosophy and Phenomenological Research 69*(2), 383–396.

Hájek, A. (2003). What conditional probability could not be. *Synthese 137*(3), 273–323.

Hawthorne, J. and M. Lasonen-Aarnio (2009). Knowledge and objective chance. In P. Greenough and D. Pritchard (Eds.), *Williamson on Knowledge*, pp. 92–108. Oxford University Press.

Leslie, J. (1992). Time and the anthropic principle. *Mind 101*(403), 521–540.

Lewis, D. (1979). Attitudes *de dicto* and *de se*. *Philosophical Review 88*(4), 513–543.

Lewis, D. (2001). Sleeping beauty: reply to Elga. *Analysis 61*(3), 171–176.

Lewis, D. K. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in Inductive Logic and Probability, Volume II*, pp. 263–293. Berkeley: University of California Press.

Magidor, O. (2015). The myth of the *de se*. *Philosophical Perspectives 29*(1), 249–283.

Manley, D. (2014). On being a random sample. Manuscript available on Manley's website.

Meacham, C. J. G. (2010). Two mistakes regarding the principal principle. *British Journal for the Philosophy of Science 61*(2), 407–431.

Moss, S. (2015). Time-slice epistemology and action under indeterminacy. In T. S. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology*, pp. 172–94. Oxford University Press.

Titelbaum, M. (2016). Self-locating credences. In A. Hájek and C. Hitchcock (Eds.), *The Oxford Handbook of Probability and Philosophy*. Oxford University Press.

Weatherson, B. (2005). Should we respond to Evil with indifference? *Philosophy and Phenomenological Research 70*(3), 613–635.