Belief Revision Revised

Joshua Edward Pearson
pearsonj@mit.edu

Abstract

I outline a novel counterexample to the principle of belief revision, Anticipation: if both learning *e* and learning *not-e* would render belief in *p* unjustified, you cannot now be justified in believing *p*. If I'm right, not only is the leading theory of belief revision false, so are various recently proposed weakenings. I develop and defend a new theory that correctly predicts the failures of Anticipation I argue for, predicated on the simple idea that one is justified in ruling out possibility just in case that possibility is sufficiently improbable.

Contents

Introduction

Belief revision theory concerns the relationship between what one is justified in believing and what one would be justified in believing were one to learn new information. One extremely plausible idea can be illustrated as follows.

> **Cookies.** The good news: your colleague has informed you that he'll be bringing homemade cookies into the department this morning. The bad news: you will not be in the department until lunch time, and you realise the following. Were you to learn that the cookies contain dairy, you would not be justified in believing you'll be eating a cookie at lunch time — you're currently trying to follow a vegan diet. And were you to learn that the cookies are dairy-free, you would not be justified in believing you'll be eating one at lunch time — there are many vegan graduate students who would have likely eaten them all by then.

Can you, nevertheless, *now* be justified in believing that you'll be eating a cookie at lunch time? Presumably not. There is a proposition *e*—*the cookies contain dairy*—such that no matter whether you were to learn it or its negation, you'd fail to be justified in believing you'll be eating a cookie at lunch time. Plausibly, this equivalent effect of learning *e* or of learning not-*e* ought to be *anticipated*, meaning you are not now justified in believing you'll be eating a cookie at lunch time.

This idea is codified by the principle of belief revision Anticipation, the claim that, roughly, if both learning *e* and learning *not-e* would render belief in *p* unjustified, you cannot now be justified in believing *p*. Anticipation is extremely plausible and is widely endorsed. Despite this, I'll argue in this paper that it's false.

My arguments raise two challenges. First, no prominent theory of belief revision can accommodate my counterexamples to Anticipation, including the dominant theory 'AGM' (Alchourrón, Gärdenfors, and Makinson 1985), as well as by various recently proposed weakenings, such as those in (Lin and Kelly 2012), (Leitgeb 2017), (Goldstein and Hawthorne 2021) and (Goodman and Salow forthcoming).[1] This naturally raises the question of whether *any* plausible theory of belief revision can accommodate my examples. I answer positively, outlining a novel theory of belief revision predicted on simple

[1]Though the situation with Goodman and Salow's (forthcoming) theory is more complex — see fn. 17 and appendix A.

idea that one is justified in ruling out a possibility just in case that possibility is sufficiently improbable.

Second, as we'll see in §1, failures of Anticipation generate problems for popular ideas about the role belief, such as its relation to rational action, assertion, and indicative conditionals. For example, if we assume that beliefs play a significant role in guiding rational action, cases in which Anticipation fails are arguably also cases in which it can be rational to avoid free evidence. If Anticipation fails, we'll need to respond to these problems or else give up on these popular ideas. I investigate both options in my conclusion.

Here's the plan. §1 outlines and motivates Anticipation in more detail. §2 presents my arguments against it. §3.1 outlines how my arguments cause trouble for present theories of belief revision, focusing on the theory given by Lin and Kelly (2012). §3.2 and §3.3 then outline my novel theory that can predict the failures of Anticipation I argue for. §4 concludes.

In keeping with the literature, I'll be making two assumptions.[2] First, binary all-out belief is coherent notion, worth theorising about, and cannot be entirely reduced to credence. Second, justified beliefs are closed under deduction: if one has justification to believe premises $P_1, ...P_n$, which mutually entail Q, then one has justification to believe Q. These assumptions rule out simple Lockeanism, the view that you have justification to believe P iff the probability that P given your evidence is sufficiently high. You might read this paper as a reductio of these assumptions — that is a debate for another time.

1 Anticipation

Here's a more precise statement of Anticipation:[3]

> **Anticipation**
> If one would not be justified in believing p were one to learn that e as total information, and one would not be justified in believing p were one to learn not-e as total information, one cannot *now* be justified in believing p.

[2](Alchourrón, Gärdenfors, and Makinson 1985), (Lin and Kelly 2012), (Leitgeb 2017), (Goldstein and Hawthorne 2021), (Hong 2023) and (Goodman and Salow forthcoming).

[3]I take the name from a previous draft of (Goodman and Salow 2023), who later discuss a generalisation of it they call 'Π−' (forthcoming). (Kraus, Lehmann, and Magidor 1990) and (Freund and Lehmann 1996) call Anticipation 'Negation Rationality'.

Why accept ANTICIPATION? We've already seen that Anticipation has considerable intuitive appeal, making plausible predictions in cases like **Cookies**. Beyond this, we can give at least two further motivations by appealing to plausible ideas about the role of justified belief.[4]

First, if ANTICIPATION were false, then strikingly infelicitous assertions would be licensed, given two plausible ideas: (i) one who would not be justified in believing p were one to learn e (as total information) is accordingly not justified in believing, and rather should doubt, the conditional *If e, p*;[5] and (ii) one is epistemically permitted to assert those propositions one is justified in believing.[6] Suppose that ANTICIPATION fails in **Cookies**: you are now justified in believing you'll eat a cookie at lunch time, even though you wouldn't be both were you to learn the cookies contain dairy and were you to learn the cookies don't contain dairy. Then you'll be in a position to assert the highly infelicitous: "I'm not sure whether I'll be eating a cookie at lunch time they contains dairy. I'm also not sure whether I'll be eating a cookie at lunch time if they don't. Nevertheless, I'll be eating a cookie at lunch time!"

Second, if ANTICIPATION is false, serious doubts emerge concerning the plausible thesis that one should, if given the opportunity, always look at free evidence before making a decision. Although this idea has not gone unquestioned, counterexamples to it have so far required agents that are risk-averse, as in (Buchak 2010), or agents that fail to know what their evidence is, as in (Salow and Ahmed 2019). The falsity of ANTICIPATION puts pressure on this claim even without assuming that rational agents can be risk-averse or ignorant of their own evidence, so long as we endorse the popular idea that justified beliefs can be used as premises in practical reasoning.[7] Suppose

[4]The notion of justification at issue is what is referred to as 'propositional' justification, rather than 'doxastic' justification. See (Silva and Oliveira 2024) for recent discussion. However, I often use the locution 'justified in believing' rather than 'have justification to believe' due to naturalness. Accordingly, I will be assuming that the epistemic agents at issue form all and only the beliefs they have justification to, in a way that is sufficient for those beliefs to be justified.

[5]We may be tempted by an even stronger principle: one has justification to believe the conditional *If e, p* iff one would have justification to believe p were they to learn e as total information. I don't make this stronger assumptions for two reasons. First, justified belief in a conditional *If e, p* can come apart from what one would believe were one to learn e in cases where e is not learnable. Second, making this stronger assumption requires care concerning triviality results (Gärdenfors 1986); see e.g. contextualist replies to these issues in (Lindström 1996), (Bacon 2015) and (Mandelkern and Khoo 2019).

[6]E.g. (Lackey 2008). (Williamson 2000) also arguably accepts this view, so long as he is interpreted as thinking that one is justified in believing all and only those propositions one knows, a position he appears sympathetic to in (Williamson forthcoming).

[7]See, for example, (Hawthorne and Stanley 2008), (Fantl and McGrath 2009) and

you justifiably believe it won't rain at your BBQ tomorrow. At the same time, it somehow turns out that checking the weather report, no matter what it says, would defeat your justification for believing it won't rain. Should you check the weather report? It's hard to see why. Checking it may cause you take take costly actions, such as cancelling your BBQ. But since you are justified in believing, and therefore can reason from, the premise that it will not rain tomorrow, such a costly action looks completely unnecessary. So you'd better not check the weather report.

So, Anticipation is highly plausible and well-motivated. Despite this, I'll be arguing that it's false. Before giving that argument, however, let me set aside a different kind of counterexample to Anticipation many have raised in conversation that I do not think works:

> **Marmite.** You justifiably believe you'll never learn whether you like Marmite. However, were you to learn that you like Marmite (say, by tasting it), you'd be justified in believing that you've learned whether you like Marmite, and were you to learn that you *don't* like Marmite (say, by tasting it), you'd again be justified in believing that you've learned whether you like Marmite.

There is a proposition e—*you like Marmite*—such that no matter whether you learn it or its negation, you'd fail to be justified in believing proposition p—*you'll never learn whether you like Marmite*. Since you are *now* nevertheless justified in believing p, don't we have a counterexample to Anticipation?

No. This argument has not paid attention to what your *total* information is in the relevant cases. When you learn you like Marmite by tasting it, you don't just learn e, you also learn that you've learned e. So your total information entails not just e, but moreover *e and you've learned whether e*. Suppose that's your total information in this case.[8]Then Anticipation only fails if, moreover, you would lose your justification for p — *you'll never learn whether you like Marmite* — were you to learn *not-(e and you've learned whether e)* as total information. But learning this proposition is of course consistent with p, making it far from clear why learning this would defeat your justification for believing p, as a counterexample to Anticipation would

(Comesaña 2020) for proponents of this idea. It is also endorsed by the work I am primarily engaging with here, such as in (Leitgeb 2017) and (Kelly and Lin 2021).

[8]Similar considerations apply even if we assume your evidence is stronger than that.

require.[9][10]

2 Against Anticipation

My argument against Anticipation comes in two steps. In §2.1, I defend previously discussed counterexamples to the stronger principle of belief revision, Preservation. In §2.2 I argue that those who find these counterexamples to Preservation persuasive should also find the counterexamples to Anticipation I outline persuasive. I respond to objections in §2.3.

2.1 Preservation Failure

Preservation is a consequence of the dominant theory of belief revision AGM (1985) and centerpiece of Leitgeb's (2014; 2017) theory:

> **Preservation**
> If one is justified in believing *p* and one is justified in leaving *e* open, then one would still be justified in believing *p* were one to learn that *e* as total information.

Here, 'one leaves *e* open' if and only if one does not believe not-*e*. Preservation possesses some intuitive plausibility. For instance, if learning that the cookies contain dairy would defeat your justification for believing you'll be eating one at lunch time, then, plausibly, you can only be justified in believing you'll eat a cookie at lunch time if you are also justified in believing the cookies don't contain dairy.

Nevertheless, Preservation faces persuasive counterexamples. I will mainly focus on the following, simple case, deployed in (Dorr, Goodman,

[9] A similar problem concerns propositions with modal operators like *might p*. You may be justified in believing *might p and might not-p*, yet this belief must be given up were you to learn *p* and were you to learn *not-p*. If, as Kratzer (2012) influentially argues, contextualist views about such modals are correct, then this is no genuine counterexample, as the expression 'might *p* and might not-*p*' expresses different propositions before and after one learns *p*. If this kind of contextualist view is incorrect, Anticipation may indeed need to be revised so as to exclude propositions with modal operators like 'might'.

[10] A final counterexample to set aside. Suppose Dr. Evil ensures that, regardless of whether I'll learn *e* or learn not-*e*, I'll have my memories which justify my belief that *p* erased, then we have a counterexample to Anticipation as stated. Strictly speaking, I should add a clause to Anticipation, and the other principles of belief revision that follow, which specifies that you do not lose any evidence when you learn *e*. I drop these qualifications for brevity.

and Hawthorne 2014) to argue against the KK principle, but that has since been used to argue against Preservation by (Stalnaker 2019, ch. 8), (Goodman and Salow 2021) and (Goodman and Salow 2023):[11]

> **Flipping for Heads.** In front of us is a fair coin, which I am going to flip it until it lands heads or has otherwise been flipped 1,000 times. You know all of this. Once I am done, I will have produced a sequence of tails, followed by a heads, or a sequence of 1,000 tails.

Plausibly, you are justified in believing that the coin won't be flipped all 1,000 times.[12] Since you know that the coin will be flipped 1,000 times if it lands on tails the first 999 times, it follows that:

(i) You are justified in believing the coin will not land tails 999 times.

At the same time, you are clearly *not* in a position to rule out the coin landing tails on the first flip, that is:

(ii) You are not justified in believing the coin won't land on tails once.

Presumably, you are also not justified in believing the coin won't land tails twice, three times, and more. However, since you *are*, by (i), justified in ruling out the coin landing tails 999 times, we will eventually reach some number of tails, k, such that while you are not justified in ruling out the coin landing tails k times, you are justified in ruling out the coin landing tails $k+1$ times. For concreteness, let's suppose that k is equal to 20. (You may suspect that the value of k is vague — I discuss this in §2.3.) Hence:

(iii) While you are not justified in believing that the coin will not land on tails 20 times, you are justified believing that the coin will not land on tails 21 times.

[11] Stalnaker (2019, ch. 8) remains sympathetic to Preservation, but does not offer a full account of how to deal with counterexamples of this kind. A similar example is also used by (Hall 1999) in relation to the surprise exam paradox.

[12] See **Objection 1** in §2.3 below for more discussion. In short, I follow much of the literature in assuming that this claim is motivated not merely by the fact that getting 1,000 tails in a row is extremely unlikely. After all, as mentioned in the introduction, I am following the literature in assuming that justified beliefs are closed under deduction, and this idea is inconsistent with the claim that high-enough probability is sufficient for justified belief (though see (2014)) . Rather, I hold that this claim gains further motivation via anti-skeptical considerations, as do e.g. (Dorr, Goodman, and Hawthorne 2014) and (Goodman and Salow 2018). See (Smith 2018) for dissent.

Here's the problem. Suppose that, in fact, you see the first 20 flips all land tails. What should you now believe? Preservation implies that, because you left it open that the first 20 flips would land all tails, you are still justified in believing the coin will not land tails 21 times in a row. This implies the absurd conclusion that you can thereby justifiably believe the next flip will land on heads. That cannot be right — absent inadmissible information (Lewis 1980), one cannot be justified in believing that a coin is both fair and will land heads when next flipped.

Moreover, there are problems for Preservation even supposing you see just the *first* flip land tails. Notice the following consequence of Preservation in this case. Since you left it open that the first flip would land on tails (as in (ii)), Preservation says that learning this does not effect your justification for believing the coin will land on tails at most 20 times (as in (iii)). So, since you're still justified in believing that the coin will land tails at most 20 times, and since you've now seen the first flip land tails, it follows that you are justified in believing there will be at most 19 *more* tails (in addition to the tails that have already occurred). Compare this to the situation before the first flip. You were then justified in believing there will be at most 20 tails, but because the first flip hadn't happened yet, you were justified in believing there will be at most 20 more tails (in addition to the tails that have already occurred).

Before seeing the first flip, you believed there will be at most 20 more tails. After seeing tails on the first flip, Preservation would have you believe there will only be at most 19 more.[13] This shift is problematic. If your knowledge that the coin is fair remains intact after seeing tails on the first flip, you have just as much reason to believe there will be 20 more tails at the start of the experiment as you do after seeing the first flip. Accordingly, it seems you should continue to believe there may be 20 more tails after seeing the first flip. If so, then after seeing the first flip you should leave open the possibility that the coin will land tails, in total, 21 times. And that means, contra Preservation, that you must revise your initial belief that the coin will land tails at most 20 times in total.

Consider an analogy. Suppose that before the coin-flipping begins, I tell you that I flipped the coin once before you entered the room and it landed on tails. This information would give you no reason to change your beliefs about what sequence the coin might produce from your present moment.

[13]Note that this fact is not itself a counterexample to Preservation. Preservation is not meant to constrain beliefs of an 'indexical' or 'de se' nature such as those concerning how many *more* times the coin will land tails, whose truth is dependent both upon what world one is in and what time one is at.

To reason as Preservation recommends in this analogous case would be to think that, because the coin landed on tails before you entered, a heads is guaranteed to come up sooner than you initially thought. This smacks of the gambler's fallacy: learning that a fair coin has landed tails should not in anyway make you think that a heads is now somehow more "overdue" than you thought previously.

Preservation therefore seems to go wrong in licensing inferences that look like the gambler's fallacy.[14] Instead, the following pattern of belief revision looks intuitively plausible: since the prospect of 21 tails, after seeing the first flip, is just as likely as the initially prospect of 20 tails, 21 tails should not be ruled out on the same grounds that 20 tails was initially not ruled out. Roughly, this means your beliefs about how many tails there might be should shift by one from those in (iii) after seeing the first flip land tails: you should now leave open the coin landing tails 21 times and rule out it landing tails 22 times.

2.2 Anticipation Failure

If we found the above argument against Preservation convincing, the following example should also convince us that Anticipation is false.

> **Flipping for Both.** In front us is a fair coin. I am going to flip it until it lands on heads at least once and on tails at least once, or until I have otherwise flipped it 1,000 times. You know all of this. Once I am done, I will have produced either a sequence of heads followed by a tails, a sequence of tails followed by a heads, a sequence of 1,000 repeating heads, or a sequence of 1,000 repeating tails.

Consider first your beliefs about how many tails in a row might be produced. Plausibly, the same considerations that applied in **Flipping for Heads** also apply here: you are justified in believing the coin won't land tails 999 times in a row, but not in believing the coin won't land on tails once, two times in a row, and so on. We will therefore eventually reach some k such that, while you can rule out $k + 1$ tails in a row, you must leave open the possibility of k tails in a row. Assuming again for concreteness that k is equal to 20, this again gives us:

[14] See (Hawthorne 2021) who uses similar intuitions surrounding the gambler's fallacy to argue for theses concerning the epistemic use of 'ought'.

(iii) While you are not justified in believing that the coin will not land on tails 20 times, you are justified believing that the coin will not land on tails 21 times.

In **Flipping for Both**, what you can believe regarding the potential number of tails in a row ought to be completely symmetric to what you can believe regarding the potential number of heads in a row. So, by running through an anaologus argument but for heads rather than for tails, we can further derive:

(iv) While you are not justified in believing that the coin will not land on heads 20 times, you are justified believing that the coin will not land on heads 21 times.

But from your justified beliefs specified in (iii) and (iv), you can derive that the coin will not land the same way, either heads or tails, 21 times in a row. Hence:

(v) You are justified in believing that the coin will not land the same way 21 times.

We now have all the required materials to argue against ANTICIPATION.

First, consider what would happen were you to learn that the first flip has landed tails. This case does not seem importantly different to the analogous scenario in **Flipping for Heads**. By (iii), you are initially justified in believing that the coin will not land tails 21 times. Yet, maintaining this belief upon seeing the first flip land tails would again be to follow a pattern of belief objectionably similar to the gambler's fallacy, for the same reasons as outlined in §2.1: you would now think a heads will come up within the next 19 heads, and so is more "overdue" than you thought before. Instead, it seems you should now leave it open that the coin will lands tails (in total) 21 times (and so 20 more times from your present moment) This, in turn, means revising your belief outlined in (v) that the coin will not land the same way 21 times.

Second, consider what would happen were you to learn that the first flip has landed *heads*. Since this case is symmetric to the case in which the first coin land tails, symmetric conclusions apply: learning that the first flip has landed heads should result in you leaving open that the coin will land heads 21 times, which means revising your belief outlined in (v) that that the coin will not land the same way 21 times.

Hence ANTICIPATION is false. Initially, you are justified in believing that the coin will not land the same way 21 times in a row. But you would not be justified in believing this were you to learn that the first flip has landed tails, and you would not be justified in believing this were you to learn that the first flip has landed heads.

This is extremely surprising. Still, my argument ought to convince anyone who is convinced by the argument against Preservation in §1. For my argument does not bring with it new substantive commitments not already present in the argument against Preservation. I have simply applied the same intuitive considerations that tell against Preservation in **Flipping for Heads** to a more complex case, **Flipping for Both**, and observed that here, these same considerations also tell against Anticipation.

Still, given Anticipation's considerable plausibility, I suspect many readers will remain cautious. So, I'll now reply to objections.

2.3 Objections

Objection 1: Lotteries

Objection: "Your arguments go wrong at the very first step: one is not permitted to believe the coin will eventually land heads. To do so is to form a belief analogous to a belief that your lottery ticket will be a loser."

Reply: Perhaps when it comes to coins, lotteries, and other cases with salient chancy-features, this objection can work. However, I worry along with (Dorr, Goodman, and Hawthorne 2014) that this kind of reply cannot be endorsed in full generality in support of Preservation and Anticipation without leading a wide-reaching skepticism. For many of our ordinary beliefs that we take to be justified arguably have a structure sufficiently similar to the above coin-flipping cases and so generate problems for Preservation and Anticipation as well.

Consider the following case from (Hall 1999). Suppose it's January 1st. Plausibly, you are justified in believing it will rain at some point this month — let's suppose your strongest justified belief is that it will rain at some point before January 15th. At the same time, you are not justified in believing that it won't rain on the 2nd. If Preservation were correct, then in general, on learning on the 2nd that it still hasn't rained, you'd still be justified in believing that it will rain at some point before January 15th. But, though this may happen in some cases, it can't be true in general. Although this case is less clean than **Flipping for Heads** — there's often not going to be probabilistic independence between the weather on different days — it's still plausible that, in at least some version of this case, learning that it hasn't rained on the 2nd should make you think that perhaps there will be an extra day without rain than you initially thought, meaning you're now only justified in believing it will rain at some point before January 16th.[15]

It's not too hard to see how to extend this example into an argument against Anticipation following a similar strategy to that in §2.2. Just consider, in addition, your strongest justified belief about how many days *with* rain there might be from January 1st — suppose it's that it will fail to rain on at least some day before January 8th. If you learn tomorrow that it has rained, you again arguably — in at least some cases — ought to extend this prediction by one. But now we have a failure of Anticipation: no matter what you learn about the weather on the 2nd, you'll have to give up your belief that it will rain on some day before January 15th and will fail to rain on some day before January 8th.

There is a general formula here. We are often justified in believing some process will eventually produce a certain output O. Yet, as time proceeds, we may remain equally justified in our beliefs regarding how quickly O will occur *from our present moment*. It is exactly cases with this structure that causes trouble for Preservation and Anticipation, in the ways outlined above. But since these beliefs are common place—you believe that not every paper you grade in this next batch will be a C; that not every second-hand item you order from eBay will be faulty; that your partner will be home from work at some point over the next two hours; that at least one kernel in this bag of popcorn will remain unpopped; etc—denying that they are justified leads to skepticism.

Objection 2: Vagueness

Objection: "Your arguments objectionably exploits assumptions that are not plausible once we accept there's vagueness about the boundaries of one's beliefs. For instance, in **Flipping for Heads**, there is no *precise* k for which k is the smallest number that you are justified in believing there will not be k heads in a row. The boundary is imprecise."

Reply: Perhaps that's right. But I have a hard time seeing how to leverage this observation in support of Preservation and Anticipation. The fact that k is vague gets the result that the above counterexamples are harder to *identify* than I have claimed. But there's a gulf from that conclusion to the further conclusion that Preservation and Anticipation are, nevertheless, *true*. For instance, consider the dominant approach to vagueness, Supervaluationism (Fine 1975), on which a claim is true iff it is true on every single admissible

[15]Compare Goodman and Salow's discussion of similar cases (Goodman and Salow 2023, p. 135, p. 137).

precisification of its vague terms. In order for Preservation (for example) to be true, it will have to hold under every single precisification of the vague term 'justifiably believes'. But the idea that one's beliefs should not follow patterns objectionably similar to the gambler's fallacy in a case like **Flipping for Heads** remains just as compelling, if not more, for precise belief states as it does for our own, fuzzy belief states.

It might be that one prefers a theory of vagueness that, unlike Supervaluationism, rejects principles of classical logic like the law of excluded middle. One will then be inclined to reject implicit premises in my argument such as that, for any number k, one either is or isn't justified in believing that there will not be k tails in a row. I take it that an approach like this is a fairly radical one. I have no new objections to it, but it is at least worth looking at alternative approaches that are consistent with e.g. the law of excluded middle, such as endorsing a theory of belief revision that rejects Preservation and Anticipation — as I'll do in §3.

Objection 3: One Philosopher's Modus Ponens...

Objection: "I agree with your conditional claim: if we accept the argument against Preservation, we should accept your argument against Anticipation. But I apply *modus tollens* where you apply *modus ponens*: we should reject the argument against Preservation."

Reply: I have some sympathy here. After all, it is not completely obvious that we should prioritize endorsing the verdicts I have argued for in **Flipping for Heads/Both** over an endorsement Anticipation. Perhaps, for theoretical reasons, it's better to deny those verdicts and hold onto Anticipation.

My problem with this reply concerns where it leaves Preservation. I doubt that a full endorsement of Preservation is a viable option. For even if we reject the argument against Preservation which generated the problem for Anticipation, recall that I initially gave a quicker argument against Preservation — this being the observation that, were you to learn the first 20 flips all land tails, you should not, as Preservation recommends, believe the next flip will land heads. I still find this objection extremely compelling, even if I accept that the case of seeing just the first flip land tails is consistent with Preservation. So, there remains pressure to give Preservation up even if one does not endorse the specific argument against that lead to the counterexample to Anticipation.

My issue from this point is that I cannot conceive of any well motivated theory of belief revision that will deny Preservation in the case in which we see 20 tails, but agree with Preservation's prediction in the case where you see just the first flip land tails. It's difficult to make my point here in complete generality. But I can at least run through very one natural attempt of constructing such a view and show why it fails. ***(Readers not concerned with the details here may skip to §3.)***

Here it that attempt. One might think that justified beliefs must be *stable* in the following sense:

> **Stability**
> If one is justified in believing p, but one would not be justified in believing p were one to learn e as total information, then one is justified in taking e to be sufficiently unlikely.[16]

And if Stability is right, perhaps we can tease apart the two purported counterexamples to Preservation. On the one hand, 20 tails in a row is extremely improbable, and so it is perfectly consistent with Stability that learning this will defeat one's justification for believing there will be at most 20 tails in a row. On the other hand, tails on the first flip is significantly likely, and so it is *inconsistent* with Stability that learning this will defeat one's justification for believing there will be at most 20 tails in a row. So Stability appears to supply an attractive view.

These appearances are misleading. In fact, Stability must be given up once we accept that 20 tails in a row can defeat one's justification for believing there will be at most 20 tails in a row. Here's why. There must be some smallest number n such that, on learning the coin as landed on tails n times in a row, you justification for believing it will not land tails 21 times in a row is defeated. Plausibly, n is not equal to 20, for learning that there have been 19 tails would presumably also defeat your belief that there will not be 21 tails in a row. But the exact value on n does not matter. Whatever its value, it will follow that learning the coin has landed tails $n - 1$ times will not defeat your belief that it will not land tails 21 times in a row. Now suppose you learn that, in fact, the coin has landed tails $n - 1$ times in a row. What will happen if you then learn, in addition, that the next flip has also landed on tails? By our stipulations, that should defeat your belief that it will not land tails 21 times in a row. But from your perspective after seeing $n - 1$ tails, it is 50% likely that it will land tails one more time. Hence we have a counterexample to Stability.

To summarize, a full endorsement of Preservation leads to intolerable problems, and it's hard to envisage viable views which avoid these prob-

[16]See (Leitgeb 2017) for an extended discussion and defence of this general idea.

lems without also endorsing the counterexample that leads to ANTICIPATION failure. The best view on offer, then, is mine: we should reject ANTICIPATION.

3 Theories of Belief Revision

I have argued that anyone who rejects PRESERVATION due to counterexamples like **Flipping for Heads** ought also reject ANTICIPATION due to counterexamples like **Flipping for Both**. This raises a challenge. No prominent theory of belief revision can accommodate my counterexamples to ANTICIPATION, including the dominant theory 'AGM' (Alchourrón, Gärdenfors, and Makinson 1985), and various recently proposed weakenings, such as those in (Lin and Kelly 2012), (Leitgeb 2017), (Goldstein and Hawthorne 2021) and (Goodman and Salow forthcoming).[17] Can *any* theory of belief revision can accommodate my examples? If the answer is "no", this may be good abductive reason to doubt my argument against ANTICIPATION. So, my aim in this section is to answer this question positively. I'll outline a novel theory of belief revision that can predict the failures of ANTICIPATION I argue for. The theory is predicated on simple and natural idea that one is justified in ruling out a possibility just in case that possibility is sufficiently improbable.

I'll begin in §3.1 by considering Lin and Kelly's (2012; 2021) theory of belief revision and diagnose why it fails to predict my counterexamples to ANTICIPATION. Doing so is instructive: my diagnosis as to where Lin and Kelly's theory goes wrong will, in part, inform how to construct an alternative. In §3.2, I outline a simple model of **Flipping for Both** which predicts that ANTICIPATION fails. I then develop this simple model into a theory in §3.3.

3.1 Lin and Kelly's Theory

Lin and Kelly's basic idea is that there's a ranking of worlds, and you're justified in believing p iff p is true throughout the top-ranked worlds (Lin and Kelly 2012; Kelly and Lin 2021). Let's call this a 'normality' ranking,

[17] The interaction between my arguments and Goodman and Salow's work is more subtle than with the other theories I have cited. For every other theory of belief revision cited, my arguments cause trouble in the simple sense that those theories fully endorse ANTICIPATION. Goodman and Salow, on the other hand, *deny* ANTICIPATION. However, my arguments here still cause trouble for them since, although Goodman and Salow outline *other* counterexamples to ANTICIPATION, their theory cannot account for the specific counterexample to ANTICIPATION I outline here. I discuss this is in detail in Appendix A.

so the highest-ranked worlds are the 'most normal' ones. For our purposes, treat 'normal' as a term of art: all that matters is whether the theory makes plausible predictions about justified beliefs, not about what's 'normal'.[18] How is normality determined? By probability. Lin and Kelly say that w_1 is more normal than w_2 (for you) iff your evidence makes w_1 sufficiently more probable than w_2. The 'most normal' worlds—the ones your beliefs leave open—are those for which no other world is more normal.[19] Summarizing, Lin and Kelly endorse:

> **Normal Belief**
> You're justified in believing p just in case, given your evidence, p is true throughout the most normal worlds.

and

> **Comparative Normality**
> World w_1 is more normal than w_2 for you just in case, given your evidence, w_1 is sufficiently more probable than w_2.

Here's a toy example. Jack is looking for his keys. His keys are either in his pocket (world w_p), in his car (world w_c), or they have been stolen (world w_s). Given Jack's evidence, w_p has a probability of $\frac{5}{10}$, w_c a probability of $\frac{3}{10}$, and world w_s a probability of $\frac{2}{10}$. Supposing that one world is sufficiently more probable than another just in case it is at least twice as likely, while w_p is sufficiently more probable than w_s (as $\frac{5}{10}$ is more than double $\frac{2}{10}$), w_p is not sufficiently more probable than w_c and w_c is not sufficiently more probable than w_s. Hence w_p and w_c are the most normal worlds, as they are not sufficiently less probable than any other world. w_s is sufficiently less probable than w_p, and so is not among the most normal worlds. So Jack is justified in believing a proposition just in case it is true in both w_p and w_c. In other words, Jack's strongest justified belief is that his keys are either in his pocket or in his car. See Figure 1.

Let's see how this approach interacts with the examples from §2. Interestingly, it provides an attractive model of **Flipping for Heads** that predicts

[18] Lin and Kelly instead use the term 'plausible'; I follow Goodman and Salow (2023; forthcoming) in using the term 'normal'. Goodman and Salow (2023, pp. 97-8) do not see it purely as a term of art, but closer to how Lewis (1973) uses the term 'similarity' in his analysis of counterfactuals, in that the term 'normal' can be useful for fixing intuitions, but that nevertheless judgments about what one is justified in believing can override judgments

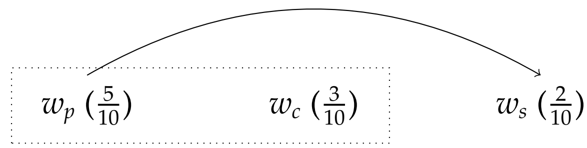

Figure 1. *Conventions.* The probability (on Jack's evidence) of a given world is written in parentheses next to it. Worlds within the dotted box are the most normal. An arrow from w to world w' indicates w is more normal than w'. (Here and below I'll only depict those arrows necessary to indicate where the dotted box should be drawn.)

the failures of Preservation argued for in §2.1.[20]

To illustrate this, we first need a set of worlds. We'll assume that each sequence the coin might produce in **Flipping for Heads** corresponds to a different possible world. In particular, let t^0 be the world in which the coin lands on heads immediately, t^1 be the world in which the coin lands on tails once before landing on heads, t^2 be the world in which the coin lands on tails twice before landing on heads, so on and so forth.

Next, the probabilities. Plausibly, these should conform to the objective chances, since in **Flipping for Heads** you know the coin is fair. Hence t^0 will initially have a probability of $\frac{1}{2}$, t^1 a probability of $\frac{1}{4}$, t^2 a probability of $\frac{1}{8}$, and so on.

Finally, to determine the normality ordering, Comparative Normality requires us to set a threshold determining when one world is sufficiently more probable than another. For simplicity, assume that w is sufficiently more probable than w' just in case it is at least 16 times more likely. (This will result in your justified beliefs in **Flipping for Heads** to be stronger than is plausible, but it will illustrate the relevant structural features of Lin and Kelly's theory just as well.) Given this, t^0 to t^3 all count as among the most normal worlds, as no world is 16 times more likely than any of them. In contrast, t_4 is excluded as t_0 (with a probability of $\frac{1}{2}$) is 16 times more likely than t_4 (with a probability of $\frac{1}{32}$). By Normal Belief, your initial strongest justified belief is therefore that the coin will land on tails no more than 3

about what's normal.

[19] I have simplified Lin and Kelly's view in assuming it is possible worlds that are ranked. In fact, Lin and Kelly rank propositions that are members of the salient partition worlds, thereby endorsing a partition/question-sensitive view of belief. I reintroduce this complication in §3.3.

[20] Note that the influential theories in (Alchourrón, Gärdenfors, and Makinson 1985) and (Leitgeb 2017) entail Preservation, and so fail to make even this prediction. Hence why I've set them aside in this discussion.

times in a row. See Figure 2.

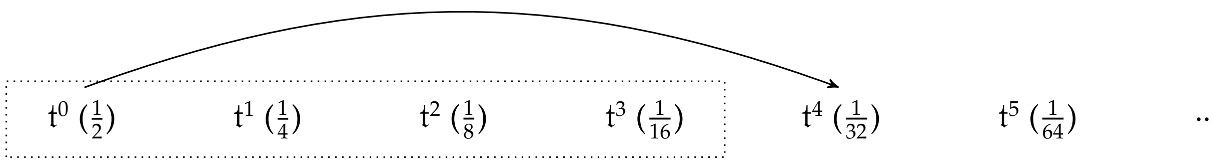

Figure 2. Flipping for Heads before any flips.

In §2.1, it was argued that Preservation fails because, upon seeing the coin flip and land tails once, your beliefs concerning how many tails there might be in a row should increase by one. On the current model, this means that given evidence excluding all and only t^0, t^4 ought to be included among the most normal worlds. That's exactly what happens. Given such evidence, t^0 has probability 0, t^1 has probability $\frac{1}{2}$ (as it is now equal to the probability that the next flip lands on heads), t^2 a probability of $\frac{1}{4}$, and so on. t^4, with an updated probability of $\frac{1}{16}$, is now among the most normal worlds since no world is at least 16 times more likely than it. See Figure 3.

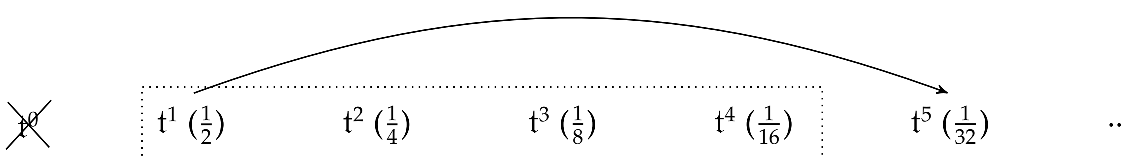

Figure 3. Flipping for Heads, after learning that the first flip landed tails. *New convention:* worlds which are crossed out are those that are incompatible with one's evidence.

This is a nice result. Indeed—on the face of it, at least—this approach vindicates the anti-gambler's-fallacy intuitions used against Preservation. So long as the coin has not yet landed on heads, you'll be justified in believing the same thing about how many *more* tails the coin may produce. In our simplified case, you'll always believe at most 3 *more* tails will occur. So far, so good! Given my arguments in §2.2, one should now expect Anticipation failures in **Flipping for Both**, too. Surprisingly, this does not happen.

In **Flipping for Both**, the coin flipping procedure no longer terminates on the first flip landing heads (but continues until it lands tails). So, to model this case, replace t^0 with worlds of the form $\mathfrak{h}^n$ — in which the coin produces n heads in a row, followed by a tails. Keeping our assumptions about the probabilities and thresholds fixed, we get the following diagram of **Flipping for Both** before the coin has been flipped:

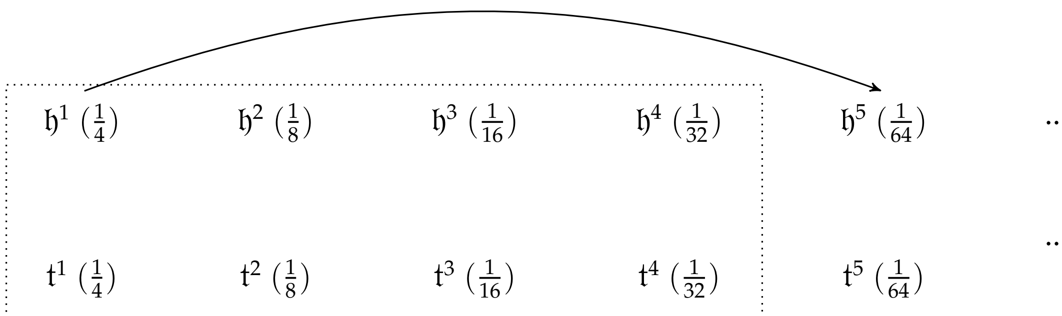

Figure 4. Flipping for Both, before any flips.

World $\mathfrak{h}^5$ and all worlds at least as improbable as it are excluded as they are at least 16 times less probable than, for example, $\mathfrak{h}^1$. All other worlds count as among the most normal. So your strongest justified belief is that there will be at most a streak of 4 heads or 4 tails in a row.[21]

Now consider what you should believe upon learning that the first flip has landed tails (symmetric considerations apply if it lands heads). This case should be no different to **Flipping for Heads**: you should revise your strongest belief about how many tails there will be in a row. Yet this is not predicted. As the first flip has landed tails, all worlds in which the first flip landed heads are now inconsistent with your evidence. Accordingly, $\mathfrak{t}^1$ increases in probability to $\frac{1}{2}$ (as it now obtains if the next flip lands heads), $\mathfrak{t}^2$ a probability of $\frac{1}{4}$, and so on. In sum, here's how the situation looks once you learn that the first flip has landed tails:

[21] Note that, surprisingly, though we have kept the orderings defined analogously to our model for **Flipping for Heads**, it was predicted in that case that one can rule out the possibility of 4 tails followed by a heads, whereas with **Flipping for Both** it is predicted that one must leave that possibility open. While my preferred approach in §3.2 will avoid this awkward consequence, I will not stake much on this here.

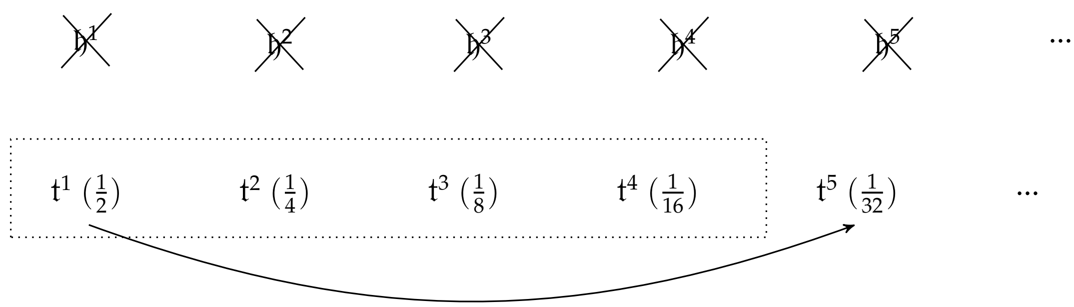

Figure 5. Flipping for Both, after first flip.

In particular, note that because t^1 is still 16 times more probable than t^5, t^5 is still excluded from the set of most normal worlds. The model therefore predicts predicts that, initially, you believe there will not be more than four tails in a row (as in Figure 4), yet upon learning that the first flip has landed tails, you *continue* to believe that there will not be four tails in a row (as in Figure 5). No counterexample to ANTICIPATION is predicted.

This is a bad result. With respect to **Flipping for Heads**, the model makes the attractive prediction that upon seeing the first flip land tails, you revise your strongest belief about how many tails in a row there will be. Indeed, to do otherwise, as I argued in §2.1, would be to change your beliefs in a way objectionably like the gambler's fallacy. But this approach fails to extend this desirable feature to **Flipping for Both**. In fact, the model *vindicates* such gambler's-fallacy-like reasoning: on seeing the first flip land tails, you should now expect the first heads to occur sooner! So, insofar as we found this approach attractive because it seemed to track our anti-gambler's-fallacy intuitions, we now see that it fails to do so once applied to **Flipping for Both**. We should therefore seek an alternative approach.

3.2 Absolute Normality — A Simple Model

It's instructive to diagnose why Lin and Kelly's approach fails to predict ANTICIPATION failure in **Flipping for Both**.[22] As I see it, the key issue is that normality is understood as a *comparative* notion: the normality of a world depends on how its probability compares to the probability of other worlds.

[22] Roughly the same diagnosis applies as to why Goodman and Salow's (2023; forthcoming) theory, along with those defended in (Goldstein and Hawthorne 2021) and (Hong 2023), also fail to predict these failures of ANTICIPATION.

For instance, we saw in Figure 4 that initially t^5 can be ruled out on the grounds that it is sufficiently less probable than t^1: t^1 has a probability of $\frac{1}{4}$ and so is 16 times more likely than t^5 which has a probability of $\frac{1}{64}$. Notice that, on learning that the first flip has landed tails, the probability of t^5 *does* significantly increase, from $\frac{1}{64}$ to $\frac{1}{32}$. But this increase in probability is not sufficient for t^5 to be included among the most normal worlds. This is because the probability of t^1 *also increases*, and importantly, it increases *at the same rate as the probability of* t^5. That is, the new probability of t^1 is $\frac{1}{2}$, and so t^1 is *still* 16 times more likely than t^5, meaning t^5 is still excluded from the set of most normal worlds.

We may therefore construct a better model by substituting this comparative notion of normality with an *absolute* notion. For instance, perhaps the normality of a world is instead determined by how its probability compares to a fixed value. Doing so could allow t^5 to become one of the most normal worlds after learning the first flip has landed on heads, since even though it is still 16 times less likely than t^1, the probability of t^5 has nevertheless substantially increased — enough so that it now counts as among the most normal.

More specifically, the alternative I am suggesting understands normality as follows:

> **Absolute Normality**
> World w counts as among the most normal for you just in case, given your evidence, the probability of w is at least τ ($0 < \tau < 1$).

While I have motivated Absolute Normality through diagnosing Lin and Kelly's approach, the picture it provides of justified belief (in conjunction with Normal Belief) is independently natural. In essence, it tells us that you're justified in believing a proposition just in case that proposition is true through all of the sufficiently probable worlds. In other words, it says that your justified in ruling a possibility out just in case that possibility is sufficiently unlikely. That is a very natural idea. Indeed, it fits nicely with the popular idea that the role of belief is to simplify reasoning by allowing agents to ignore possibilities that are sufficiently unlikely, endorsed by, for example, (Harsanyi 1985), (Lance 1995), (Lin 2013) and (Ross and Schroeder 2014).[23]

[23] At least with respect to what an agent justifiably believes at a specific time, this approach is similar to the one defended by (Levi 1967), which also predicts a threshold τ such that any possibility falling below that threshold is believed not to obtain. But there

I'll develop this idea more in §2.3. For now, let's see how we can use Absolute Normality to construct an attractive model of **Flipping for Both** that successfully predicts the failure of Anticipation argued for in §2.2. I'll illustrate this toy model using bar charts, as follows. Each bar along the x-axis will represent a different possibility in **Flipping for Both**. The height of each bar represents that possibility's probability. τ is then represented as a point on the y-axis with a dotted line running through it. Worlds with a probability of at least τ will be shaded in, representing those worlds that are most normal. Your justified in believing p just in case it is true in all of the shaded worlds.

Setting $\tau = \frac{1}{16}$, this is how things initially look in **Flipping for Both**:

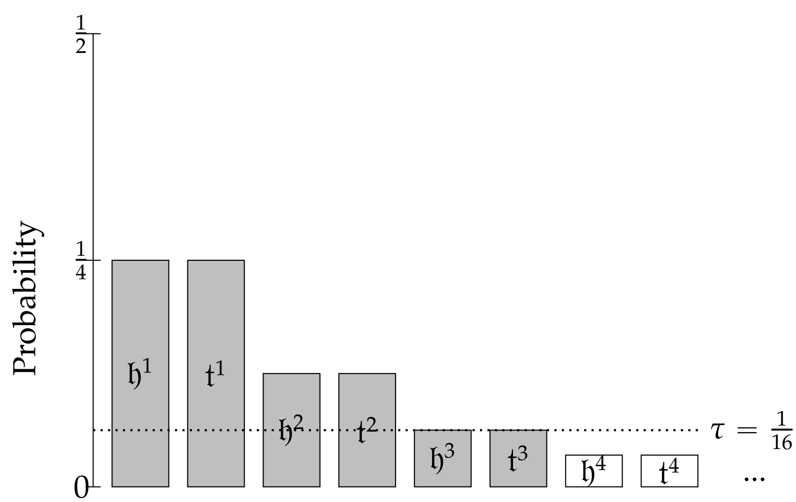

Figure 6. New model for **Flipping for Both** before any flips. *Conventions.* The height of each bar represents the probability of the world written inside the bar. Bars shaded in represent all and only those worlds consistent with what the relevant agent is justified in believing.

That is, initially you believe there will at most 3 heads in a row and at most 3 tails in a row — and so at most 3 of the same in a row. Upon learning that the

are important differences. Levi further holds that whether a proposition can be justifiably believed depends on how informative that proposition is ((Dorst and Mandelkern 2022) defend a similar idea). The way Levi measures the informativeness of a proposition p means that p can become *more* informative as information in gained and possibilities in W are ruled out. This feature means that the dynamics of Levi's theory — that is, how an agent's beliefs change across times — turns out to be quite different to the predictions to the theory I endorse here and cannot, for instance, make the desired predictions concerning **Flipping for Both**.

first flip has landed tails, the $\mathfrak{h}^n$ worlds are eliminated, and the probability of the $\mathfrak{t}^n$ worlds are adjusted, giving us the following updated diagram:

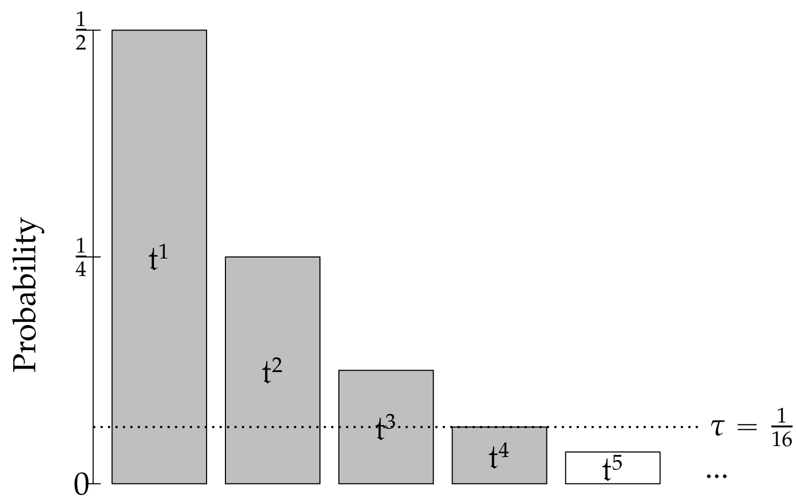

Figure 7. Alternative **Flipping for Both** model after first flip lands tails.

That is, after the first flip has landed tails, one is no longer justified in believing there will be at most 3 of the same in a row, as $\mathfrak{t}^4$ is now a possibility consistent with what one is justified in believing. A symmetric result will hold in the case in which the first flip lands on heads, giving us the ANTICIPATION failure argued for in §2: initially you believe that there will be no more than three of the same, and this belief will be revised both on learning that the first flip lands on tails and on learning that the first flip lands on heads.

3.3 Absolute Normality — Developing the View

This simple model offers us an attractive picture of what is going on in **Flipping for Both**. But do ABSOLUTE NORMALITY and NORMAL BELIEF combine to give us a generally plausible theory of belief and, in turn, belief revision? Not quite. For two reasons, the official theory I'll endorse is more complex.

The first reason is that we need a "global" threshold requirement on justified belief. For all that has been said so far, nothing prevents you from justifiably believing propositions that are highly unlikely. Consider a case with four worlds — w_1, w_2, w_3 and w_4 — with, respectively, probabilities $\frac{4}{10}$, $\frac{3}{10}$, $\frac{2}{10}$ and $\frac{1}{10}$. If τ is set at $\frac{4}{10}$, then only w_1 counts as among the most normal worlds. We'll thereby predict that the relevant agent is justified is believing that w_1 obtains, a proposition they should only take to be 40% likely.[24]

To solve this, I'll adopt the following idea, influenced by the approaches in (Goldstein and Hawthorne 2021), (Hong 2023) and (Goodman and Salow forthcoming). Here's the intuitive idea. If it is sufficiently unlikely that you're in a world that is at least τ-likely, you should concede that things are not as normal as you usually have the right to suppose. Nevertheless, it's plausible that you can still be justified in believing you're in a world that is *pretty* normal, or perhaps *somewhat* normal, so long as it is sufficiently likely that they are in a pretty normal/somewhat normal world. In essence: if the set of τ-likely worlds is itself not so likely, introduce the next-most likely world into your belief set, then the next-most likely after that, until that set is itself sufficiently likely. We'll see how to implement this intuitive idea formally in a moment.

The second reason is that we should introduce "question-sensitivity". Consider:[25]

> **Dime or Nickel.** I am about to perform the coin-flipping procedure from **Flipping for Both**. However, to decide whether I'll use a dime or a nickel for my coin, I roll a fair die: I'll use a dime if it lands even and a nickel otherwise.

Whether I use a dime or a nickel is irrelevant to how long a streak of tails I might get. Both are fair coins. So there should be no difference in your beliefs in **Flipping for Both** and in **Dime or Nickel**. However, NORMAL BELIEF and ABSOLUTE NORMALITY will not predict this if, for **Dime and Nickel**, we have to distinguish between worlds in which the same sequence is produced but by a different coin. Doing so means that, for instance, there will two worlds in which the coin lands tails 4 times, each with a probability of $\frac{1}{32}$. If the threshold τ is at $\frac{1}{16}$, then in **Dime or Nickel**, but not **Flipping for both**, you can rule out the coin landing tails 4 times in a row. This is absurd.

The most promising response here is to follow various other authors in

[24] A particularly acute version of this problem arises if *none* of the possibilities have a probability of at least τ. In that case, the set of most normal worlds is empty, implying the absurd conclusion that the relevant agent is justified in believing a contradiction. While theories of weak belief (e.g. (Dorst and Mandelkern 2022) and (Holguín 2022)) will be happy with justified beliefs in unlikely propositions, they will not be happy with justified beliefs in contradictory propositions.

[25] Thanks to Jonathan Fiat for bringing my attention to this kind of case.

this literature — Lin and Kelly (2012), Leitgeb (2017),[26] Hong (2023) and Goodman and Salow (forthcoming) — and endorse a question-sensitive account of justified belief.[27] The idea is that, rather than saying what one is justified in believing is determined by an invariant set of worlds, what one is justified in believing is rather determined in part by what *question* is salient.[28] Generally, when a distinction between two possibilities is not relevant for answering the salient question—that is, when those two possibilities provide the same answer to that question—those two possibilities will not be distinguished in that context. Applying this thought to **Dime or Nickel**, the idea will be that if the relevant question is *What sequence will be produced?* then we need not distinguish between worlds in which a different coin is used so long as those coins produce the same sequence in those worlds. So, at least with respect to that question, we'll get the desirable prediction that your beliefs in **Dime or Nickel** should be the same as in **Flipping for Both**, as your justified beliefs in each case will be formed relative to the same way of carving up the possibilities.[29][30]

Here's how to introduce both complications to construct a generally attractive account of justified belief and belief revision. Let W be the set of possible worlds, Q a partition of it, and $[\cdot]_Q$ a function taking worlds of W to the cell of Q they are a member of. Q represents the relevant question. Let P be a probability function defined over W, and let P_e be that function conditionalized upon your evidence, represented by proposition e — this will represent your subjective probabilities.[31]

[26]Leitgeb notably faces a similar problem to the one just outlined; see (Staffel 2016, pp. 1731-2).

[27]See also (Holguín 2022), (Blumberg and Lederman 2020) and (Yalcin 2018).

[28]I speak loosely here to remain neutral between semantic versions of question-sensitivity (e.g. (Goodman and Salow 2021) and (Holguín 2022)) and subject-sensitive versions (e.g. (Leitgeb 2017)).

[29]Introducing question-sensitivity opens up a potential reply to my arguments against Anticipation in §2.2: perhaps my arguments illicitly shift the relevant question. I discuss this in detail in Appendix A, section 4. In short, I argue that my arguments still go through so long as we hold fixed the question: *What sequence will be produced?*

[30]I deny this fix is ad hoc. Indeed, question-sensitivity fits nicely with the motivation given above that beliefs simplify reasoning. For partitioning a complex and large set of possibilities into fewer chunks is another way for agents to simplify their reasoning. The idea that coarse-gaining possibilities allows agents to simplify decision problems is also common in economics; see, for example: (Ahn and Ergin 2010), (Epstein, Marinacci, and Seo 2007) and (Gul, Pesendorfer, and Strzalecki 2017).

[31]I'll be assuming throughout that evidence updates monotonically, and that probabilities update by conditionalisation, as is common throughout the literature: (Lin and Kelly 2012), (Leitgeb 2014), (Goldstein and Hawthorne 2021), (Goodman and Salow forthcoming).

Given question sensitivity, we'll now reinterpret Absolute Normality to be about *answers*, rather than worlds. Specifically: you must now leave open any *answer to the relevant question* that is at least τ likely. Formally, this constraint says that the following set — the set of "sufficiently normal" worlds — must always be a subset of the strongest proposition you believe:

$$N_{Q,e} =_{def} \{w \in e : P_e([w]_Q) \geqslant \tau\}$$

However, given the global threshold constraint now introduced, we cannot equate your strongest justified belief with $N_{Q,e}$. Instead, we'll define a second threshold, T ($\frac{1}{2} < T < 1$) such that if $P_e(N_{Q,e}) < T$, the strongest proposition you believe contains more worlds than the sufficiently normal ones (i.e. those in $N_{Q,e}$).

The following theory does exactly that:[32]

Absolute Normal Belief (ANB)
Given evidence e, your strongest justified belief relative to Q, $B_{Q,e}$ is equal to the smallest set S such that:

(i) $N_{Q,e} \subseteq S$.
S contains all of the sufficiently normal worlds.

(ii) For all $w_1, w_2 \in e$, if $Pr_e([w_1]_Q) \geqslant Pr_e([w_2]_Q)$, then $w_2 \in S$ only if $w_1 \in S$.
S includes all worlds that are at least as probable as any world in S.[33]

(iii) $Pr_e(S) \geqslant T$. ($\frac{1}{2} < T < 1$)
S is sufficiently likely.

I take ANB to be an attractive alternative to the other theories considered here, one that makes better predictions in **Flipping for Both**, thereby denying Anticipation.

ANB also has various other interesting and attractive results. Since stating these results is a slightly tangential topic, I've placed a full overview them in Appendix B. Briefly, however, it is worth noting that though ANB predicts Anticipation fails, it nevertheless validates weaker principles of

[32]Condition (ii) has gained seriously traction in the literature on weak belief; Dorst and Mandelkern (2022) call it "filtering"; Holguin (2022) calls it "cogency". Conditions (ii) and (iii) are endorsed by various other theories of belief — see (Goldstein and Hawthorne 2021), (Hong 2023) and (Goodman and Salow forthcoming). It is condition (i) that makes my theory distinctive.

belief revision. Moreover, ANB can, unlike its competitors, predict various attractive restrictions of belief ”accretion” principles which dictate when beliefs can be *formed* on learning new information.

4 Life Without Anticipation

Let's take stock. I have argued that Anticipation fails in cases like **Flipping for Both**. All theories of belief revision defended thus far fail to get this result. In response, I have developed an alternative theory which can get this result, which stands up to scrutiny, and is moreover predicated on the natural idea that you are justified in ruling out possibilities that are sufficiently unlikely.

However, outlining this theory only answers one of the two challenges I set out in the introduction. The second challenge, recall, was that failures of Anticipation generate challenges to popular ideas concerning the role belief plays in other philosophically significant areas. As we saw in §1, given these popular ideas, if Anticipation is false, then bizarre and infelicitous assertions are licensed and one can be rational in avoiding free evidence. Supposing we accept my arguments against Anticipation, that leaves us with two options. We must either find some way to live with these awkward consequences, or we must deny the popular ideas about belief that were used to derive them. I'll close by examining both options.

Let's consider, first, denying those popular ideas about belief. My arguments concerning rational evidence avoidance depended on justified beliefs playing a substantive role is rational decision making. In particular, they depended on the idea that the propositions one is justified in believing can be used as premises in practical reasoning. Maybe justified beliefs play no such role. However, so long as justified beliefs play *some* important role in rational decision making — one that cannot be reduced to the role played by rational credences — those who hope to pursue this strategy need to tell us what this role is.

My arguments concerning infelicitous assertions relied on a contentious principle connecting belief revision and beliefs in conditionals. Triviality results, from e.g. (Lewis 1976) and (Gärdenfors 1986), give us reason to doubt there is any such neat connection here. But thinking there is *some* connection between belief revision and conditionals is irresistible. So, those who hope to ameliorate this awkward consequence of Anticipation failures by denying there is a straightforward connection between belief revision and beliefs in conditionals will need to tell us what the not-so-straightforward connection is.

Here's one tempting idea. The theories we've discussed distinguish justified beliefs from propositions that are part of your evidence. Maybe the roles given to the former should be reserved for the latter. For example, perhaps indicative conditionals are connected to *evidence* revision, rather than *belief* revision; and perhaps it's permissible to rely only on propositions that are part of your *evidence* in your reasoning, rather than all your justified beliefs. The problem? This saps (mere) justified beliefs of their philosophical interest—evidence has supplanted them.[34] I'm inclined to resist this conclusion, but my arguments do seem to forge a new path to reaching it.

Can we, instead, learn to live with the awkward consequences of accepting these ideas about belief alongside ANTICIPATION failure? Here's a promising avenue. Cases in which I've claimed ANTICIPATION fails are plausibly cases in which there is 'iteration failure' for belief: cases where you're justified in believing p without being justified in believing you're so justified. Suppose that in **Flipping for Both** you must leave open a streak of 19 heads/tails in a row, but you can rule out a streak of 20. If so, it will be extremely difficult for you to distinguish your actual case from one where you must instead leave open a streak 20 heads/tails in a row but can rule out 21. So, you plausibly won't be justified in believing that you're justified in believing there won't be a streak of 20. Following (Williamson forthcoming) and (Carter and Hawthorne forthcoming), we can leverage this to offer an account of where agents that make the outlined bizarre assertions or decline free evidence are going wrong. Though they are acting in accordance with their justified beliefs — in a way that is epistemically *permitted* — they are nonetheless acting in a way that is epistemically *risky*. That is, since they cannot justifiably believe they are acting in accordance with their justified beliefs, they are not in a position to justifiably believe they are acting in a way that is epistemically permitted. This approach is of course in need for further elaboration. But if it works, perhaps we can learn to live without ANTICIPATION.[35]

[34] Compare Williamson's (2000, ch. 9) influential knowledge-first approach, in which one's evidence is all and only the propositions one knows. Williamson recognizes that if there are some propositions one knows that are not part of one's evidence, then in some sense evidence, not knowledge, comes first.

[35] Thanks to Kevin Dorst, Jonathan Fiat, Jeremy Goodman, Richard Roth, Bernhard Salow, Jack Spencer, Bob Stalnaker and Roger White for various conversations that strongly influenced the ideas in this paper. I've also benefited from insightful comments from Lisa Cassell, Cal Fawell, Jessica Heine, Justin Khoo, Sebastian Liu, Philipp Mayr, Sonia Pavel, Kayleigh Rodgers, Abraham Matthew, Kieran Setiya, Eliot Watkins and Eliza Wells.

References

Ahn, David S. and Haluk Ergin (2010). "Framing Contingencies". In: *Econometrica* 78.2, pp. 655–695. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/40664487 (visited on 02/29/2024).

Alchourrón, Carlos E., Peter Gärdenfors, and David Makinson (1985). "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions". In: *Journal of Symbolic Logic* 50.2, pp. 510–530. DOI: 10.2307/2274239.

Bacon, Andrew (2015). "Stalnaker?s Thesis in Context". In: *Review of Symbolic Logic* 8.1, pp. 131–163. DOI: 10.1017/s1755020314000318.

Blumberg, Kyle and Harvey Lederman (2020). "Revisionist Reporting". In: *Philosophical Studies* 178.3, pp. 755–783. DOI: 10.1007/s11098-020-01457-4.

Buchak, Lara (2010). "Instrumental Rationality, Epistemic Rationality, and Evidence-Gathering". In: *Philosophical Perspectives* 24.1, pp. 85–120. DOI: 10.1111/j.1520-8583.2010.00186.x.

Carter, Sam and John Hawthorne (forthcoming). "Dogmatism & Inquiry". In: *Mind*.

Comesaña, Juan (2020). *Being Rational and Being Right*. Oxford, United Kingdom: Oxford University Press.

DeRose, Keith (1995). "Solving the Skeptical Problem". In: *Philosophical Review* 104.1, pp. 1–52. DOI: 10.2307/2186011.

Dorr, Cian, Jeremy Goodman, and John Hawthorne (2014). "Knowing Against the Odds". In: *Philosophical Studies* 170.2, pp. 277–287. DOI: 10.1007/s11098-013-0212-9.

Dorst, Kevin and Matthew Mandelkern (2022). "Good Guesses". In: *Philosophy and Phenomenological Research*. DOI: 10.1111/phpr.12831.

Epstein, Larry G., Massimo Marinacci, and Kyoungwon Seo (2007). "Coarse contingencies and ambiguity". eng. In: *Theoretical Economics* 2.4, pp. 355–394. ISSN: 1555-7561. URL: http://hdl.handle.net/10419/150101.

Fantl, Jeremy and Matthew McGrath (2009). *Knowledge in an Uncertain World*. Ed. by Matthew McGrath. New York: Oxford University Press.

Fine, Kit (1975). "Vagueness, Truth and Logic". In: *Synthese* 30.3-4, pp. 265–300. DOI: 10.1007/bf00485047.

Freund, Micheal and Daniel Lehmann (1996). "On Negation Rationality". In: *Journal of Logic and Computation* 6.2, pp. 263–269. DOI: 10.1093/logcom/6.2.263.

Gärdenfors, Peter (1986). "Belief Revisions and the Ramsey Test for Conditionals". In: *Philosophical Review* 95.1, pp. 81–93. DOI: 10.2307/2185133.

Goldstein, Simon and John Hawthorne (2021). "Knowledge From Multiple Experiences". In: *Philosophical Studies* 179.4, pp. 1341–1372. DOI: 10.1007/s11098-021-01710-4.

Goodman, Jeremy and Bernhard Salow (2018). "Taking a Chance on KK". In: *Philosophical Studies* 175.1, pp. 183–196. doi: 10.1007/s11098-017-0861-1.

— (June 2021). "Knowledge from Probability". In: *Electronic Proceedings in Theoretical Computer Science* 335, pp. 171–186. issn: 2075-2180. doi: 10.4204/eptcs.335.15. url: http://dx.doi.org/10.4204/EPTCS.335.15.

— (2023). "Epistemology Normalized". In: *Philosophical Review*.

— (forthcoming). "Belief Revision Normalized". In: *Journal of Philosophical Logic*.

Gul, Faruk, Wolfgang Pesendorfer, and Tomasz Strzalecki (Jan. 2017). "Coarse Competitive Equilibrium and Extreme Prices". In: *American Economic Review* 107.1, pp. 109–37. doi: 10.1257/aer.20141287. url: https://www.aeaweb.org/articles?id=10.1257/aer.20141287.

Hall, Ned (1999). "How to Set a Surprise Exam". In: *Mind* 108.432, pp. 647–703. doi: 10.1093/mind/108.432.647.

Harsanyi, John C. (1985). "Acceptance of Empirical Statements: A Bayesian Theory Without Cognitive Utilities". In: *Theory and Decision* 18.1, pp. 1–30. doi: 10.1007/bf00134075.

Hawthorne, John (Feb. 2021). "The Epistemic Use of 'Ought'". In: *Conditionals, Paradox, and Probability: Themes from the Philosophy of Dorothy Edgington*. Oxford University Press. isbn: 9780198712732. doi: 10.1093/oso/9780198712732.003.0010. eprint: https://academic.oup.com/book/0/chapter/267641856/chapter-pdf/38900980/oso-9780198712732-chapter-10.pdf. url: https://doi.org/10.1093/oso/9780198712732.003.0010.

Hawthorne, John and Jason Stanley (2008). "Knowledge and Action". In: *Journal of Philosophy* 105.10, pp. 571–590.

Holguín, Ben (2022). "Thinking, Guessing, and Believing". In: *Philosophers' Imprint*, pp. 1–34.

Hong, Frank (2023). "Prefaces, Knowledge, and Questions". In: *Ergo: An Open Access Journal of Philosophy* 10.n/a. doi: 10.3998/ergo.4631.

Kelly, Kevin and Hanti Lin (2021). "Beliefs, Probabilities, and Their Coherent Correspondence". In: *Lotteries, Knowledge, and Rational Belief: Essays on the Lottery Paradox*. Ed. by IgorEditor Douven. Cambridge University Press, pp. 185–222. doi: 10.1017/9781108379755.011.

Kratzer, Angelika (Jan. 2012). *Modals and Conditionals: New and Revised Perspectives*. Oxford University Press. isbn: 9780199234684. doi: 10.1093/acprof:oso/9780199234684.001.0001. url: https://doi.org/10.1093/acprof:oso/9780199234684.001.0001.

Kraus, Sarit, Daniel Lehmann, and Menachem Magidor (1990). "Nonmonotonic Reasoning, Preferential Models and Cumulative Logics". In: *Artificial Intelligence* 44.1-2, pp. 167–207. doi: 10.1016/0004-3702(90)90101-5.

Lackey, Jennifer (2008). *Learning From Words: Testimony as a Source of Knowledge*. Oxford: Oxford University Press.
Lance, Mark Norris (1995). "Subjective Probability and Acceptance". In: *Philosophical Studies* 77.1, pp. 147–179. DOI: 10.1007/bf00996316.
Leitgeb, Hannes (2014). "The Stability Theory of Belief". In: *Philosophical Review* 123.2, pp. 131–171. DOI: 10.1215/00318108-2400575.
— (2017). *The Stability of Belief: How Rational Belief Coheres with Probability*. Oxford University Press.
Levi, Isaac (1967). *Gambling with Truth: An Essay on Induction and the Aims of Science*. Cambridge: MIT Press.
Lewis, David (1973). *Counterfactuals*. Blackwell.
— (1976). "Probabilities of Conditionals and Conditional Probabilities". In: *Philosophical Review* 85.3, pp. 297–315. DOI: 10.2307/2184279.
— (1980). "A Subjectivist's Guide to Objective Chance". In: *Studies in Inductive Logic and Probability, Volume II*. Ed. by Richard C. Jeffrey. Berkeley: University of California Press, pp. 263–293.
— (1996). "Elusive knowledge". In: *Australasian journal of Philosophy* 74.4, pp. 549–567.
Lin, Hanti (2013). "Foundations of Everyday Practical Reasoning". In: *Journal of Philosophical Logic* 42.6, pp. 831–862. DOI: 10.1007/s10992-013-9296-0.
Lin, Hanti and Kevin Kelly (2012). "Propositional Reasoning That Tracks Probabilistic Reasoning". In: *Journal of Philosophical Logic* 41.6, pp. 957–981. DOI: 10.1007/s10992-012-9237-3.
Lindström, Sten (1996). "The Ramsey Test and the Indexicality of Conditionals: A Proposed Resolution of Gärdenfors' Paradox". In: *Logic, Action and Information*. Ed. by André Fuhrmann and Hans Rott. de Gruyter.
Mandelkern, Matthew and Justin Khoo (2019). "Against Preservation". In: *Analysis* 79.3, pp. 424–436. DOI: 10.1093/analys/any051.
Ross, Jacob and Mark Schroeder (2014). "Belief, Credence, and Pragmatic Encroachment". In: *Philosophy and Phenomenological Research* 88.2, pp. 259–288. DOI: 10.1111/j.1933-1592.2011.00552.x.
Salow, Bernhard and Arif Ahmed (2019). "Don't Look Now". In: *British Journal for the Philosophy of Science* 70.2, pp. 327–350. DOI: 10.1093/bjps/axx047.
Shear, Ted and Branden Fitelson (2018). "Two Approaches to Belief Revision". In: *Erkenntnis* 84.3, pp. 487–518. DOI: 10.1007/s10670-017-9968-1.
Silva, Paul and Luis R. G. Oliveira (2024). "Propositional Justification and Doxastic Justification". In: *Routledge Handbook of the Philosophy Evidence*. Ed. by Maria Lasonen-Aarnio and Clayton M. Littlejohn. Routledge.

Smith, Martin (2018). “Coin Trials”. In: *Canadian Journal of Philosophy* 48.5, pp. 726–741. DOI: 10.1080/00455091.2017.1381936.
Staffel, Julia (2016). “Beliefs, Buses and Lotteries: Why Rational Belief Can’t Be Stably High Credence”. In: *Philosophical Studies* 173.7, pp. 1721–1734. DOI: 10.1007/s11098-015-0574-2.
Stalnaker, Robert (2019). *Knowledge and Conditionals: Essays on the Structure of Inquiry*. Oxford University Press.
Williamson, Timothy (2000). *Knowledge and its Limits*. Oxford University Press.
— (forthcoming). “Justifications, Excuses, and Sceptical Scenarios”. In: *The New Evil Demon*. Ed. by Fabian Dorsch and Julien Dutant. Oxford: Oxford University Press.
Yalcin, Seth (2018). “Belief as Question-Sensitive”. In: *Philosophy and Phenomenological Research* 97.1, pp. 23–47. DOI: 10.1111/phpr.12330.

Appendix A: Goodman and Salow's Theory

Though much of what I say about Lin and Kelly's (2012) theory can also be said for Goodman and Salow's (forthcoming), Goodman and Salow's theory raises three subtle issues that need discussing. A.1 briefly outlines their view. A.2 discusses the fact that their can can predict Anticipation failure in cases where one gains evidence that cross-cuts the relevant question, and A.3 discusses the fact that on an extension of their approach — on which the notion of a 'de se' question is introduced — they predict even more failures of Anticipation. However, I'll argue in both cases that they still do not have the resources to make the desired predictions in **Flipping for Both**. A.4 discusses an alternative strategy in which the counterexample to Anticipation in **Flipping for Both** is explained away, rather than vindicated, by appeal to shifting contexts. I argue that it's unsatisfactory.

A.1 Goodman and Salow's Theory

We can understand Goodman and Salow's theory as replacing condition (i) from Absolute Normal Belief with the simpler condition that $B_{Q,e}$ contains the worlds belonging to the mostly likely answer to Q.[36] I'll presuppose that we are keeping a single question, Q, fixed, and so will not explicitly parameterize beliefs to the relevant question.

GandS
Given evidence e, your strongest justified belief, B_e is equal to the smallest set S such that:

(i) S contains all worlds in the cell of Q that P_e says is most likely.

(ii) For all $w_1, w_2 \in e$, if $Pr_e([w_1]_Q) \geqslant Pr_e([w_2]_Q)$, then $w_2 \in S$ only if $w_1 \in S$.

(iii) $Pr_e(S) \geqslant T$. ($\frac{1}{2} < T < 1$)

Setting $T = \frac{15}{16}$, it's easy to check that, on the face of it, GandS will make exactly the same predictions in **Flipping for Heads** and **Flipping for Both** as Lin and Kelly's theory does, as illustrated in Figures 2, 3, 4 and 5.

[36]See Goodman and Salow (forthcoming, §6), who also provide other equivalent statements of their theory.

A.2 Cross-cutting Questions

As Goodman and Salow (forthcoming, §8) note GandS predicts failures of Anticipation in cases where your evidence cross-cuts the relevant question. They consider:

> **Celebrity Hike.** 101 celebrities go on a hike in Runyon Canyon. A paparazzo shadowing them notices a hiking pole on the trail. On inspection, he notices some fingerprints. He knows that Michael Jackson and Beyoncé were on the hike, and that Michael always hikes wearing one glove on his right hand and Beyoncé always hikes wearing one glove on her left hand. After inspecting the pole further, the paparazzo discovers whether the fingerprints were made by a left or right hand.

They outline the following model. Let W consist of 200 worlds, one for each hand that might have made the fingerprints on the pole: one for Michael's left hand, one for Beyoncé's right hand, and one for either hand of the remaining 98 celebrities. Let P be a uniform probability distribution over W, and Q the question *Who dropped the pole?* Setting $T = 0.99$, it follows that before looking fingerprints, the paparazzo is justified in believing p: *someone other than Michael or Beyoncé dropped the pole*. (Each celebrity other than Michael or Beyoncé has $\frac{1}{100}$ probability of dropping the pole, as they have two potential hands which could have dropped it; Michael and Beyoncé each only have a $\frac{1}{200}$ probability.) However, were the paparazzo to learn that the fingerprints on the pole were made by a left hand, or were he to learn the fingerprints were made by a right hand, the paparazzo must, contra Anticipation, give up his belief in p, as then all of the non-eliminated possible answers—which will include just one of *Michael dropped the pole* or *Beyoncé dropped the pole*—will now be equally likely. (Given it's a left hand, Micheal is just as likely as anyone else; given it's a right hand, Beyoncé is just as likely as anyone else.)

Nevertheless, GandS still cannot offer a plausible model of **Flipping for Both**, which is not naturally modeled as involving cross-cutting evidence. The relevant anti-Anticipation intuitions are solicited just by considering the question *What sequence of heads/tails will the coin produce*. Further, note that it's difficult to model Anticipation failure in **Flipping for Both** using GandS *even if* we model the case as one in which your evidence cross-cuts the relevant question. As outlined in the following footnote, it's possible to generate models that look *close* to the Anticipation failure argued for in §2.2, but these models have bizarre skeptical consequences and otherwise require the use of highly gerrymandered partitions of W as the relevant question.[37]

A.3 De Se Questions

Goodman and Salow describe the following further counterexample to Anticipation (2021, Appendix C) (forthcoming, §10):

> **Flipping for All Heads.** A coin flipper will simultaneously flip 100 fair coins until they all simultaneously land heads. Then he will flip no more.

Plausibly, one in justified in believing both that it will take at least a few simultaneous flips before they land on all heads, and also that the coins will all land on heads simultaneously *eventually*. The strongest proposition one is justified in believing is therefore that the number of trials required before the coin flipping ends falls in some interval $[n, m]$, with $n > 1$.

If that's right, Anticipation fails. Consider the proposition p: *there will be between 2 and m simultaneous flips*. One is initially justified in believing this as it is entailed by one's strongest justified belief. However, were one to learn that some coins landed tails on the first simultaneous flip, then, by soliciting similar anti-gambler's-fallacy intuition invoked in §2.1, it seems one should give up believing p as one should now take it to be possible that there will be $m + 1$ trials. At the same time, were one to learn that every coin landed heads on the first simultaneous flip, one should now think that the process took exactly 1 trial, therefore also giving up one's belief that p. Anticipation fails.

Modeling this case requires Goodman and Salow to introduce a further tool: *de se* questions. Skipping on the formal details—see (Goodman and Salow 2021, Appendix C) and (Goodman and Salow forthcoming, §10)—here's

[37] For instance, where W is as described in §2.2 for **Flipping for Both**, let:

$$Q = \{\mathit{For}\ 1 \leqslant n \leqslant 4 : \{\mathfrak{h}^n, \mathfrak{t}^n\}\} \cup \{\mathit{For}\ n \geqslant 5,\ x \in \{\mathfrak{h}, \mathfrak{t}\} : \{x^n, x^{n+1}, ..., x^{1,000}\}\}$$

That is, the possible answers are: *1 heads/tails in a row,..., 4 heads/tails in a row*, *More than 4 heads in a row*; *More than 4 tails in a row*. Setting $T = \frac{15}{16}$, GandS entails that one believes, before the first flip, that there will not be more than 4 heads/tails in a row, yet this belief must be given up regardless of whether one learns that the first flip landed heads or that it landed tails. This model is extremely gerrymandered and moreover has the skeptical implication that, on learning that the first flip landed heads (tails), one must now take *any* number of heads (tails) to be possible.

the basic idea. Goodman and Salow can model this kind of ANTICIPATION failure as occurring when the salient question is *When will the coins all land heads together*, where this has the three answers (to use a simple example) *extremely soon*; *an extremely long time from now*; or *in between*. Unlike with regular questions, which worlds belong to which members of this de se partition changes depending on the moment of time you're in. For instance, although the world in which the process takes $1,000,000$ attempts may initially count as *an extremely long time from now*, after you've observed $999,999$ attempts all fail to produce 100 simultaneous heads, the world in which is takes $1,000,000$ attempts now counts as *extremely soon*. Selecting an appropriate threshold, you'll continue to believe the answer *in between* as your strongest belief for so long as the coin-flipping continues. However, ANTICIPATION can fail since the propositional content of *in between* changes over time: at first *in between* may have propositional content $[n, m]$, but after the first flip has failed to land all heads, it will have content $[n+1, m+1]$.

But we still lack the resources to model ANTICIPATION failure in **Flipping for Both**. There is no natural *de se* question for which ANTICIPATION failure in **Flipping for Both** is predicted. For example, a natural de se question to consider here is: *How long a streak of heads/tails, will the coin produce from now?*. Answers to this question, such as *one more*, will have differing propositional content as the coin flipping in **Flipping for Both** unfolds: before any flip it will have content $\{\mathfrak{h}^1, \mathfrak{t}^1\}$, but after, say, the first flip lands heads, it will have content $\{\mathfrak{h}^1\}$. Setting $T = \frac{15}{16}$, GANDS predicts that before the first flip one believes that that there will be no more than in total four heads in a row, but also that this belief is preserved after learning the first flip has landed heads, meaning ANTICIPATION failure is not predicted.

A.4 Appeals to Question-Shifting

Goodman and Salow — and, potentially, Lin and Kelly — may hope to utilize their framework in a different way. Perhaps they can use question-sensitivity to *explain away*, rather than predict, my counterexample to ANTICIPATION, in the same way early contextualists about knowledge used shifts of context to explain away apparent failures of single-premise closure (DeRose 1995) (Lewis 1996).

Before outlining the details, I'll start with an objection. Given that Goodman and Salow predict *other* counterexamples to ANTICIPATION, trying to explain away the counterexample from **Flipping for Both** looks ad hoc. This is especially so given that their **Flipping for All Heads** counterexample exploits intuitions similar to those grounding my **Flipping for Both** counterexample.[38]

Nevertheless, here's how an appeal to question-sensitivity might help. (I recommend the reader first re-familiarizes themselves with my argument against Anticipation in §2.2.) We can distinguish between three questions:

Q_{Tails}: *How many consecutive tails will there be at the beginning of the sequence?*

Q_{Heads}: *How many consecutive heads will there be at the beginning of the sequence?*

Q_{Same}: *How long will the opening consecutive sequence, either of heads or tails, be?*

Distinguishing between these questions may allow us to offer a debunking explanation of my argument against Anticipation in §2.2. Crucially, my argument relied on inferring (v) from (iii) and (iv): (see p. 9, substituting in 3 for 20 and 4 for 21):

(iii) While you are not justified in believing that the coin will not land on tails 3 times, you are justified believing that the coin will not land on tails 4 times.

(iv) While you are not justified in believing that the coin will not land on heads 3 times, you are justified believing that the coin will not land on heads 4 times.

(v) You are justified in believing that the coin will not land the same way 4 times.

However, according to GandS, there is no *single* question for which (v) can be legitimately derived. Setting $T = \frac{15}{16}$, we can see that (iii) is true with respect to Q_{Tails} — you can rule out 4 tails but not 3 — but (iv) is not: with respect to Q_{Tails}, sequences that start with a heads all belong to the same cell, meaning you leave any number of heads open. Symmetrically, (iv) is true with respect to Q_{Heads} but (iii) is not. As for Q_{Same}, *neither* (iii) nor (iv) hold: you can neither rule 4 tails nor 4 heads out. My argument therefore appears insensitive to subtle changes of the relevant question, illicitly deriving (v) from premises that are only true with respect to different questions.

I take this to be the most promising reply Goodman and Salow, and perhaps Lin and Kelly too (though they discuss question-sensitivity less

[38]Note that, in contrast, ANB will be able to predict the Anticipation failure in **Flipping for All Heads** if also extended to allow for de se questions.

extensively than Goodman and Salow), can make towards my argument against ANTICIPATION. Nevertheless, even ignoring the extent to which it feels ad hoc, I think it's unsatisfactory.

The problem is that the above three questions are not the only ones that might be relevant. We might instead be asking the question Q_{Seq}: *What sequence will the coin produce*, which is naturally interpreted as a trivial partition on W, consisting of a singleton-cell for each member of W. With respect to Q_{Seq}, GANDS makes the same predictions to those in §3.1 (Figures 3 and 4). It therefore faces the same objection: it vindicates, rather the prevents, the objectionable gambler's-fallacy-like patterns of belief revision. Setting the threshold $T = \frac{15}{16}$, you initially believe in **Flipping for Both** that t^4 might but t^5 won't occur. After seeing the first flip land tails, you'll *still* believe that t^4 might but t^5 won't occur. So you'll consider a heads to be more overdue after seeing the first flip land on tails. So, insofar as these kinds of worries were motivating our departure from PRESERVATION and ANTICIPATION, this question-sensitive debunking approach should strike us as unsatisfactory.[39]

[39]Note further that Q_{Same} can only be a relevant question if we allow for evidence that cross-cuts the salient question. That means we'd have to deny "ORTHOGONALITY", discussed in Appendix B — an issue I want to remain neutral on for this paper.

Appendix B: Consequences of Absolute Normal Belief

I'll start by, for convenience, restating the theory defended in the main paper. First, a few stipulations and terminological clarifications:

- Let W be the set of possible worlds, Q a partition of it, and $[\cdot]_Q$ a function taking worlds of W to the cell of Q they are a member of. Q represents the relevant question. Propositions, denoted by e.g. 'p', 'q' and 'e', are subsets of W.

- For simplicity, I am going to assume that the relevant agent — which I'll refer to using second-person pronouns like 'you' — initially has a trivial body of evidence equivalent to W. Moreover, when you learn a new proposition p as total information, your new evidence becomes your old evidence intersected with p.[40]

- Let P be a probability function defined over W—representing your initial subjective probabilities—and let P_p be that function conditionalized on proposition p.[41]

- Where τ is a member of the unit interval $[0, 1]$, $N_{Q,p}$ is the set of of "sufficiently normal" worlds, relative to question Q and proposition p:

$$N_{Q,p} =_{def} \{w \in p : P_p([w]_Q) \geqslant \tau\}$$

The theory defended in the main paper is then as follows:

Absolute Normal Belief (ANB)
Given evidence e, your strongest justified belief relative to Q, $B_{Q,e}$ is equal to the smallest set S such that:

(i) $N_{Q,e} \subseteq S$.
S contains all of the sufficiently normal worlds.

(ii) For all $w_1, w_2 \in e$, if $Pr_e([w_1]_Q) \geqslant Pr_e([w_2]_Q)$, then $w_2 \in S$ only if

[40] This is a widespread assumption in the literature I am engaging with — e.g. (Lin and Kelly 2012), (Leitgeb 2014), (Goldstein and Hawthorne 2021) and (Goodman and Salow forthcoming).

[41] $P_p(q) = \frac{P(p\&q)}{P(p)}$ when $P(p) > 0$ and is undefined otherwise.

$w_1 \in S$.
S includes all worlds that are at least as probable (relative to Q) as any world in S.

(iii) $Pr_e(S) \geqslant T$. $(\frac{1}{2} < T < 1)$
S is sufficiently likely.

I'll separate the consequences of ANB to be outlined into two categories: first, its further implications regarding belief revision (B.1) — principles constraining relationship when you should/shouldn't *give up* a belief on learning new information — and its consequences regarding belief "accretion" (B.2) — principles constraining when you should/shouldn't *form* a belief on learning new information.

Going forward, I'll hold fixed the relevant question Q. Most principles of belief revision/accretion will fail if the relevant question is allowed to shift. So, strictly speaking, we are only interested in version of these principles where "believes" is always interpreted as being relativized to a single question, Q. Due to this, I can drop the explicit Q parameter for e.g. $B_{Q,e}$ and $N_{Q,e}$. Further, when I am considering your initial beliefs — in which you evidence is equal to W — I will also drop the evidence parameter.

B.1 ANB and Belief Revision

We saw in §3.2 how condition (i) enables ANB to predict failures of two key principles of belief revision:

Preservation
If one is justified in believing p and one is justified in leaving e open, then one would still be justified in believing p were one to learn that e as total information.

Anticipation
If one would not be justified in believing p were one to learn that e as total information, and one would not be justified in believing p were one to learn not-e as total information, one cannot *now* be justified in believing p.

However, the theory of belief revision ANB leaves us with is far from completely unconstrained. Consider the following natural weakening of Anticipation:

Reversal Anticipation
If you would be justified in believing not-*p* were you to learn *e* as total information, and you would be justified in believing not-*p* were you to learn not-*e* as total information, you're not now justified in believing *p*

Reversal Anticipation only rules out you believing *p* in cases where learning learning *e*/not-*e* would justify belief in *not-p*; unlike Anticipation, it allows belief in *p* if learning *e*/not-*e* would merely prevent you from justifiably believing *p*. Reversal Anticipation is even more plausible than Anticipation, and ANB entails it:

Upshot 1. ANB entails Reversal Anticipation.

Proof. Assume you believe $\neg p$ on learning *e* and believe $\neg p$ on learning $\neg e$. Then, by condition (iii) of **ANB**, and since $T > \frac{1}{2}$, it follows that:

$(*)$ $Pr_e(\neg p) > \frac{1}{2}$ and $Pr_{\neg e}(\neg p) > \frac{1}{2}$.

By the law of total probability and $(*)$, $Pr(\neg p) = Pr_e(\neg p)Pr(e) + Pr_{\neg e}(\neg p)Pr(\neg e) > \frac{1}{2}Pr(e) + \frac{1}{2}Pr(\neg e) = \frac{1}{2}$. So $P(\neg p) > \frac{1}{2}$. Since $T > \frac{1}{2}$, by condition (iii) of ANB we also know that $P(B) > \frac{1}{2}$. Hence it must be that $\neg p \cap B \neq \emptyset$, giving us the desired result that you don't believe *p*. $\square$

Now consider an analogous weakening of Preservation:

Reversal Preservation
If your justified in believing *p* and you're not justified in believing $\neg e$, then you wouldn't be justified in believing $\neg p$ were you to learn *e* as total information.

In general, ANB predicts counterexamples to Reversal Preservation:

Countermodel. $Q = \{\{w_1, w_2\}, \{w_3\}\}$; $Pr(w_1) = \frac{60}{100}$, $Pr(w_2) = \frac{1}{100}$, $Pr(w_3) = \frac{39}{100}$; $\tau = \frac{40}{100}$, $T = \frac{60}{100}$. Then $B = \{w_1, w_2\}$, yet $B_{\{w_2, w_3\}} = \{w_3\}$. So you start out believing $w_1 \vee w_2$, leave open $w_2 \vee w_3$, but on learning the latter, come to believe $\neg(w_1 \vee w_3)$, contra Reversal Preservation.

However, all such countermodels must play the following trick: the proposition you learn *cross-cuts* the relevant question. It's unclear whether or not we should allow for such cases. If it's possible you'll learn something that cross-cuts Q, then distinctions which cross-cut Q are relevant, suggesting that the relevant question is something more fine-grained than Q. (Lin and Kelly 2012) effectively presuppose cross-cutting evidence cannot happen; (Goodman and Salow forthcoming) remain neutral. I'll also remain neutral here. But, suppose we follow Lin and Kelly and accept:

Orthogonality
When the relevant question is Q, you can only learn a proposition e if e is the union of members of Q. *You cannot learn propositions that are orthogonal to the relevant question.*

Then ANB predicts Reversal Preservation. That's because Orthogonality implies the following more general constraint:[42]

Orthogonality*
Relative to Q, e is a learnable proposition only if for all w_1, w_2 in W: $\frac{Pr([w_1]_Q)}{Pr([w_2]_Q)} = \frac{Pr_e([w_1]_Q)}{Pr_e([w_2]_Q)}$ when w_1, w_2 are members of e and $Pr([w_2]_Q) > 0$.

Upshot 2. ANB and Orthogonality entail Reversal Preservation.

> *Proof.* Assume you believe p, meaning $B \subseteq p$, and assume you don't believe $\neg e$, meaning $e \cap B \neq \varnothing$. Consider $v \in e \cap B$. v is also in p, as $B \subseteq p$. Since you believe p, by condition (ii) of ANB we know that, for any $w \in \neg p$, $P([v]_Q) > P([w]_Q)$ (otherwise B would contain $\neg p$-worlds). By Orthogonality*, we thereby also know that $P_e([v]_Q) > P_e([w]_Q)$. It follows by condition (ii) of ANB that a $\neg p$-world is a member of B_e only if v is a member of B_e. Since since $v \in p$, we get the desired result that you do not believe $\neg p$ on learning e. □

In sum: ANB does not leave us with a completely unconstrained theory of belief revision, getting us, for example, Reversal Anticipation. And given

[42]See (Goodman and Salow forthcoming, §6) for discussion. Note that endorsing Orthogonality would rule out **Celebrity Hike** as a potential counterexample to Anticipation. This may be spun as an advantage to the extent the one finds the anti-Anticipation verdict in **Celebrity Hike** counter-intuitive.

Orthogonality, it entails Reversal Preservation, too. ANB thus supplies an attractive way of giving up Anticipation without having to give up too much more.

That being said, ANB does predict counterexamples to the following principle of belief revision, whether or not we accept Orthogonality:

Cautious Monotinicity
If you're justified in believing p and you're justified in believing e, you would still be justified in believing p were you to learn e as total information.

Countermodel. $Q = \{\{w_1\}, \{w_2\}, \{w_2\}\}$, $P(w_1) = \frac{6}{10}$, $P(w_2) = P(w_3) = \frac{2}{10}$, $\tau = \frac{1}{4}$ and $T = \frac{6}{10}$. Initially, $B = w_1$, so you believe you're in w_1, and you believe the deductive conseuqence that you're either in w_1 or w_2. Yet on learning $\{w_1, w_2\}$, since $P_{\{w_1,w_2\}}(w_2) = \frac{1}{4} \geqslant \tau$, $w_2 \in B_{\{w_1,w_2\}}$ meaning, contra Cautious Monotonicity, you no longer believe you're in w_1.

These failures of Cautious Monotonicity are not particularly intuitive. However, I think that ANB has enough other advantages that these consequences can be accepted on theoretical grounds. It's also worth noting that standard Lockeanism predicts failures of Cautious Montinicty (Shear and Fitelson 2018, §2.1), as do the extensions of Goodman and Salow's theory (discussed above in §A.3) that allow for 'De Se' questions (Goodman and Salow forthcoming, §10).

B.2 ANB and Belief Accretion

What does ANB say about how beliefs are *gained* upon learning new information? Generally, ANB predicts, along with the theories in Goodman & Salow and Lin & Kelly, that both of the following principles fail:[43]

[43] Another relevant principle is "Proof by Cases", which, roughly, states that one must believe p if one would believe p were one to learn e and one would believe p were one to learn not-e. As Goodman and Salow note (forthcoming, §7, fn. 27), Proof by Cases is equivalent to Fronloading under natural assumptions, so I'll only consider the latter here. (Note that they consider a principle they call '$\Pi+$' — a generalization of Proof by Cases)

Frontloading
You're justified in believing p on learning e as total information only if you're is initially justified in believing $(\neg e \vee q)$.

Inductive Conservatism
If you're justified in believing e but not p, then you would still not be justified in believing p were you to learn e as total information.

ANB accommodates similar countermodels to these principles as those outlined in (Goodman and Salow forthcoming, §7).[44] Goodman and Salow concede that these countermodels are not particularly intuitive, but accept them nonetheless on theoretical grounds.

However, matters are different for ANB if we assume Orthogonality, by which it entails a natural restriction of both principles. In particular, in cases where your initial strongest belief is exactly the proposition that you're in one of the sufficiently normal worlds — so $B = N$ — both principles hold. In intuitive terms, it is only special cases where it is sufficiently unlikely that you're in one of the sufficiently normal worlds that these principles breakdown. This can be seen by proving the restricted version of Frontloading, which entails Inductive Conservatism:[45]

Upshot 3: Assuming Orthogonality, ANB entails Frontloading in the special case where $B = N$.

Proof. Assume you don't believe $(\neg e \vee p)$, meaning $(e\&\neg p) \cap B \neq \varnothing$. Let v be an element of $(e\&\neg p) \cap B$. Assume for contradiction that you believe p on learning e, meaning $B_e \subseteq p$. This means that v, a $\neg p$-world, is not an element of B_e. By condition (i) of ANB, this means:

(1) $P_e([v]_Q) < \tau$.

Since $v \in B$ and $B = N$, we also know that:

(2) $P([v]_Q) \geqslant \tau$.

[44]Note that they call Inductive Conservatism '$\Box+$'.

[45]For suppose Inductive Conservatism fails: (1) you believe e, (2) you don't believe p, yet (3) you believe p were you to learn e. If Frontloading held, (3) entails that you initially believe $\neg e \vee p$. But by (1) and (2), your initial belief set must contain some $(e\&\neg p)$-worlds.

So $[v]_Q$ must decrease in probability conditional on e. Assuming Orthogonality, this can only happen if e is inconsistent with $[v]_Q$. But we've selected v such that $v \in e$. Contradiction. Hence you do not believe p on learning e. □

Dropping Orthogonality, ANB accommodates countermodels to Frontloading, even if $B = N$:

Countermodel. $Q = \{\{w_1, w_2\}, \{w_3, w_4\}\}$, $P(w_1) = P(w_3) = P(w_4) = \frac{3}{10}$, $P(w_2) = \frac{1}{10}$, $\tau = \frac{3}{10}$ and $T = \frac{6}{10}$. Let $p = \{w_2, w_3, w_4\}$ and $q = \{w_1, w_3, w_4\}$. Initially, $B = N = W$, so you don't believe $\neg p \vee q$. However, $B_p = \{w_3, w_4\}$, so you do believe q on learning p, contrary to Frontloading.

Nevertheless, ANB still predicts a restricted version of Inductive Conservatism, even without Orthogonality.

Upshot 4 ANB entails Inductive Conservatism in the special case where $B = N$.

Proof. Assume you believe e: $B \subseteq e$. A new belief will be gained after learning e only if there exists $w \in B$ such that $w \notin B_e$. As $B = N$, and by condition (i) of ANB, this can only happen if:

($*$) There exists $w \in B$ such that: $P_e([w]_Q) < \tau$.

But recall that $B \subseteq e$. And $[w]_Q \subseteq B$. So $[w]_Q$ can only increase in probability conditional on e, meaning ($*$) cannot be satisfied. Hence no new beliefs are gained on learning e. □

In sum: although ANB is, *in general*, as permissive as Goodman and Salow's (forthcoming) theory when it comes to belief accretion, it has an advantage. In particular, condition (i) of ANB — unique to this theory — can be utilized to predict restrictions of these principles of belief accretion, which helps explain their intuitive plausibility.