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 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.

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.** Good news: your colleague has sent a department-wide email stating he has baked surplus cookies and will be bringing them into the department this morning. Unfortunately, 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. Meanwhile, 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 paper has been modified to better serve as a writing sample for job applications. The original version is available on my website: `joshuaedwardpearson.com`.

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 ANTICIPATION is, in fact, 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), (Hong 2023) and (Goodman and Salow 2023; 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 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. But note that rejecting these assumptions would itself be a significant result.

[1] Though the situation with Goodman and Salow's theory is a little more complex than with the others I have cited, as I explain in fn. 17.

[2] Both assumptions are made, at least implicitly, across the literature I engaging with, e.g. in (Alchourrón, Gärdenfors, and Makinson 1985), (Lin and Kelly 2012), (Leitgeb 2017), (Lin 2019), (Goldstein and Hawthorne 2021), (Hong 2023) and (Goodman and Salow 2023).

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*.

Three clarifications. First, I follow the literature on belief revision in stating these principles in the subjunctive mood: they concern what one *would* believe *were* one to learn new information.[4] This phrasing invites counterexamples to Anticipation not usually of interest to those studying belief revision. For example, if 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: I'm justified in believing *p*, but wouldn't be were I to learn *e* or were I to learn not-*e*. For that reason, it's perhaps better to understand such principles in terms of *conditional belief*: whether, on hypothetically adding *e* to your body of evidence, you'd be justified in believing *p*.[5] However, since this technical notion of conditional belief is less familiar, and since the subjunctive gloss will serve well enough for our purposes, I shall stick with it, setting problem cases of this kind aside.

Second, although I will often informally drop it, the qualification that the principle only concerns propositions that are learned "as total information" is crucial for avoiding further counterexamples. Consider:

> **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.

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

[4] E.g: (Gärdenfors 1986), (Huber 2013), (Leitgeb 2017), (Lin 2019) and (Goodman and Salow forthcoming).

[5] See (Ramsey 1926, p. 247) who famously uses this notion of conditional belief in his discussion of indicative conditionals.

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. The case in which you learn you like Marmite by tasting it is a case in which you learn more than just *e*, you also learn a stronger proposition: *you've learned whether e*. Let's suppose your total information in this case is characterised by the proposition e'. To properly assess whether we have a counterexample to ANTICIPATION, the further case to consider is not one in which you learn *not-e*, but rather a case in which you learn as total information *not-*e'. Since e' entails that you've learned whether you like Marmite, *not-*e' is compatible with scenarios in which you *haven't* learned whether you like Marmite. So it is not at all clear that learning *not-*e' gives you reason to give up your belief in *p*, meaning we have no counterexample to ANTICIPATION.[6]

Third, the notion of justification at issue is what is referred to as 'propositional' justification, rather than 'doxastic' justification.[7] 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.

Why accept ANTICIPATION? As we have seen, ANTICIPATION has considerable intuitive appeal, and it makes plausible predictions in simple cases like **Cookies**. Beyond this, we can give two further motivations by examining how failures of ANTICIPATION interact with other plausible ideas about belief. First, failures of ANTICIPATION license bizarre assertions, 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*';[8] and (ii) one is epistemically permitted to assert those

[6]A similar purported counterexample concerns propositions with modal operators such as *might p*. For example, 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 a different propositions before and after one learns *p* or *not-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'.

[7]See (Silva and Oliveira 2024) for recent discussion.

[8]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 won't make this stronger assumption here; doing so requires care concerning triviality results (Gärdenfors 1986). I am sympathetic to contextualist replies to these triviality concerns, see (Lindström 1996), (Bacon 2015) and (Mandelkern and Khoo 2019).

propositions one is justified in believing.[9] If both ideas are right, then failures of Anticipation license bizarre assertions. Suppose that Anticipation fails in **Cookies**: you are now justified 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 if it contains dairy. I'm also not sure whether I'll be eating a cookie at lunch time if it doesn't. Nevertheless, I'll be eating a cookie at lunch time!"

Second, if Anticipation is false, serious doubts emerge concerning the 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.[10] Suppose 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.[11]

In sum, Anticipation is a highly plausible and well-motivated principle. However, as I'll argue in the next section, it's also false.

[9]See, for instance, (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).

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

[11]This challenge is similar to Kripke's second dogmatism puzzle (Kripke 2011). However, the problem presented by Anticipation failure is harder than Kripke's puzzle—solutions to the latter can't obviously be applied to the former. For example, Carter and Hawthorne (forthcoming) propose to solve Kripke's puzzle by noting that (substituting "knows" for "justifiably believes"), although looking at the evidence may risk losing a justified belief that p, it may also result in obtaining a higher-order justified belief that one justifiably believes p. However, if Anticipation fails, one is *guaranteed* to lose the justified belief in question if one looks at the evidence, meaning Carter and Hawthorne's solution cannot be applied here. The same applies to Salow's recent response to Kripke's second dogmatism, see (Salow forthcoming, fn. 8).

2 Against Anticipation

My argument against Anticipation comes in two steps. I begin in §2.1 by outlining known and widely accepted counterexamples to a strictly 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, too. I respond to objections to my arguments in §2.3.

2.1 Preservation Failure

Let's begin, then, with Preservation—a consequence of the dominant theory of belief revision AGM (1985) and centrepiece of Leitgeb's (2014; 2017) recent 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 decisive counterexamples. I will mainly focus on the following, simple case. It was deployed in (Dorr, Goodman, and Hawthorne 2014) to argue against the KK principle, but has since been used to argue against Preservation in (Stalnaker 2019, ch. 8), (Goodman and Salow 2021) and (Goodman and Salow 2023):[12]

> **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.

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

Plausibly, you are justified in believing that the coin won't be flipped all 1,000 times. 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.

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 fair 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 here. 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 *now*. 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. And 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] That is, it seems in tension with the following rough principle:

No Gambler's Fallacy
If agent A_1 at t_1 is in an epistemically equivalent position as to whether coin C_1 will produce sequence x_1 as agent A_2 at t_2 is in as to whether coin C_2 will produce sequence x_2, then A_1 is justified in believing C_1 will not produce x_1 iff A_2 is justified in believing C_2 will not produce x_2.[15]

[13]Note that this fact is not itself a counterexample to Preservation. Preservation is not meant to constrain beliefs of an 'indexica'l 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.

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

[15]Much will depend on how we flesh out the notion of "epistemically equivalent positions". The intuitive notion will do for the purposes of this paper. Were I to be more precise, I would say that they are in epistemically equivalent positions just in case, with respect to each agent's respective evidence e_1 and e_2, the probability that C_1 produces x_1 conditional on e_1 is equal to the probability that C_2 produces x_2 conditional on e_2.

Flipping for Heads illustrates that PRESERVATION is inconsistent with NO GAMBLER'S FALLACY. Let you-before-the-first flip be A_1 at t_1, and you-after-the-first-flip be A_2 at t_2. A_2 knows that the first flip has landed tails, but that does not seem to matter — if A_1 cannot rule out their coin producing 20 *more* tails, neither can A_2. Yet, if PRESERVATION is right, A_2, unlike A_1, can justifiably believe the coin will not land on tails 20 more times in a row.

NO GAMBLER'S FALLACY is extremely plausible. If it's right then, contra PRESERVATION, upon seeing the first flip lands tails in **Flipping for Heads**, your beliefs about how many tails there might be in (iii) should shift by one: 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** apply also in this slightly more complex case: 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:

(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, to maintain this belief in the face of seeing the first flip land tails would objectionably violate NO GAMBLER'S FALLACY, for the same reasons as outlined in §2.1. Instead, you should now leave it open that the coin will lands tails 21 times. 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 the coin will not land the same way 21 times.

And now we have a counterexample to ANTICIPATION. 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. So ANTICIPATION is false.

This is extremely surprising. That being so, this argument against ANTICIPATION ought to be completely convincing to anyone who is convinced by the argument against PRESERVATION in §1. For this 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 considerations also tell against ANTICIPATION.

Still, given ANTICIPATION's considerable plausibility, I recognise that many readers will remain suspicious of my arguments. So, before moving on, I'll reply to various 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 is on the money. 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, 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 that can't generally be true — plausibly, there are versions of this case where this prediction must shift by one and you are now only justified in believing it will rain at some point before January 16th.

It's not too hard to see how to extend this example 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 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 can 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 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 exploit 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 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 'believes'. But the idea that one's beliefs should not violate No 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 classical logic. One will then be inclined to reject implicit premises in my argument such as that, for any number k, one either believes or does not believe that there will not be k tails in a row (an instance of the law of excluded middle). 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 classical logic, 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 prioritise accepting No Gambler's Fallacy over Anticipation. Both have considerable intuitive appeal. Perhaps it's No Gambler's Fallacy rather than Anticipation that should go.

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 that depended on No Gambler's Fallacy, the first, quicker argument against Preservation I gave still feels compelling — 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.

So, there remains pressure to give Preservation up even if one does not endorse the argument against it based on No Gambler's Fallacy. However, my issue now is that I cannot conceive of any plausible view that will have the desired consequence of endorsing the counterexample to Preservation just mentioned yet not endorsing the counterexample that depended on No Gambler's Fallacy. 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, it looks as if we can tease apart the two purported counterexamples to Preservation as follows. 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, it is significantly likely that the next flip will land on tails, 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.

This appears to be an attractive view. Stability looks very plausible. And while this view will reject No Gambler's Fallacy, it nevertheless avoids endorsing Preservation.

But 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

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

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 summarise, a full endorsement of Preservation leads to intolerable problems, and views which avoids these problems while still resisting my arguments against Anticipation are unmotivated. 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), as well as by various recently proposed weakenings, such as those in (Lin and Kelly 2012), (Leitgeb 2017), (Goldstein and Hawthorne 2021), (Hong 2023) and (Goodman and Salow 2023; Goodman and Salow forthcoming).[17] It is therefore natural to wonder whether *any* theory of belief revision can accommodate my examples. 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 outline why it fails to predict my counterexamples to Anticipation. Doing so is instructive: my diagnosis as to where Lin and Kelly's theory

[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 the original version of this paper (Pearson ms, §3.2).

goes wrong will, in part, inform how to construct an alternative. The same basic story can be told concerning Goodman and Salow's theory (2023; forthcoming);[18] I stick to Lin and Kelly's theory as it is simpler. In §3.2 I'll outline a simple model of **Flipping for Both** which predicts that ANTICIPATION fails. I then develop this simple model into a theory of belief revision in §3.3.

3.1 Lin and Kelly's Theory

The basic idea behind Lin and Kelly's approach is as follows (Lin and Kelly 2012; Kelly and Lin 2021). On their approach, there is a ranking of possible worlds, and one is justified in believing a proposition p just in case p is true throughout the most highly ranked worlds. I'll describe a world's position in this ranking as the 'normality' of that world, so that the most highly ranked worlds are those that are 'most normal'.[19] But note that this is purely for terminological convenience. For our purposes, we care only about whether the theory that follows makes plausible predictions about what one is justified in believing; we do not care whether the theory provides a plausible analyses of our pre-theoretical concept of 'normality'.[20]

So, how do Lin and Kelly propose how the most 'normal' worlds are determined? Lin and Kelly propose that we do so by examining the *probability* of that world. In particular, they hold that for an agent S, w' is more normal than world w just in case, given S's evidence, w' is sufficiently more probable than w. w is then counted among the 'most normal' — and is therefore a world that, for all S is justified in believing, obtains — just in case no other world is more normal than it.[21]

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, we get that 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

[18]See (Pearson ms, §3.2).

[19]Lin in Kelly instead use the term 'plausible'; I follow Goodman and Salow in using the term 'normal'.

[20]See also (Goodman and Salow 2023, p. 97), who treat "normal" as a technical notion in a similar way as Lewis (1973) treats the term "similar" in his theory of counterfactuals.

[21]This overview of Lin and Kelly's is simplified in that Lin and Kelly do not rank worlds, but rather rank propositions that are members of a salient partition of the set of worlds. In other words, interpreting a partition of possible worlds as a question, they endorse a theory of justified belief that is "question-sensitive". I reintroduce this complication in §3.3.

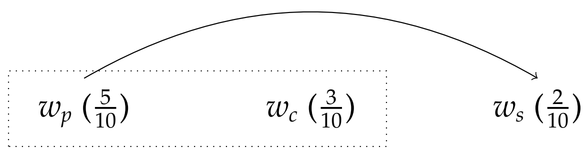

Figure 1. Conventions — the fraction next to each world represents the probability of that world given the relevant agent's (in this case, Jack's) evidence. Worlds inside the dotted box are those that are among the most normal worlds. An arrow from world w to world w' represents that w is more normal than w'. (Note the figures below do not depict all such arrows, but only those necessary to indicate where the dotted box should be drawn.)

than w_s. On Lin and Kelly's approach, this means that worlds w_p and w_c are among the most normal, as they are not sufficiently less probable than any other world. w_s, on the other hand, is sufficiently less probable than w_p, and so is not among the most normal worlds. Jack is therefore 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. Figure 1 provides an illustration.

How does Lin and Kelly's theory interact with the counterexamples to Preservation and Anticipation argued for in §2? Interestingly, Lin and Kelly's approach provides an attractive model of **Flipping for Heads** that predicts the failures of Preservation argued for in §2.1.[22]

To illustrate this, we first need a set of worlds to work with. We'll assume that each different 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, we need to assign each world a probability. Plausibly, the probability should conform to the objective chances, for we have supposed that in **Flipping for Heads** you know that the coin is fair. So, before you have seen the coin flipped at all, t^0 will have a probability of $\frac{1}{2}$, t^1 a probability of $\frac{1}{4}$, t^2 a probability of $\frac{1}{8}$, so on and so forth.

Finally, to determine what you are justified in believing (i.e. the set of most normal worlds), we need to set a threshold for when one world counts as sufficiently more probable than another. For simplicity, I am going to assume

[22]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.

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 far stronger than is plausible, but it will still serve our purposes of illustrating the structural features of Lin and Kelly's theory, such as how Preservation can fail.) 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 $\frac{1}{32}$). Your strongest justified belief is therefore that the coin will land on tails no more than 3 times in a row. Figure 2 illustrates.

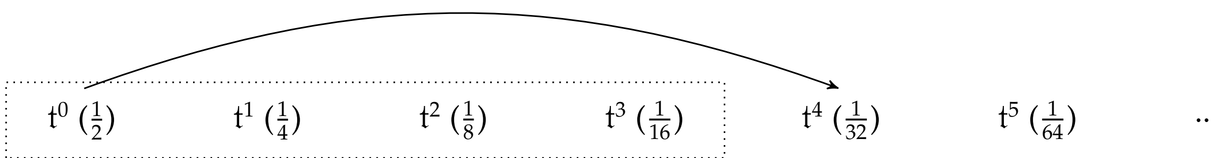

Figure 2. Flipping for Heads before any flips.

In §2.1, it was argued that Preservation fails because, upon seeing the coin flip and land tails once, your beliefs concerning for how many tails there might be in a row should increase by one. On the current model, this means that given evidence which excludes t^0, t^4 ought to be included among the most normal worlds. And that's exactly what this model predicts. Given evidence that excludes t^0, t^0 should be given a probability of 0, t^1 will now have a probability of $\frac{1}{2}$ (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. t^5, however, remains excluded as it is 16 times less likely than t^1. Figure 3 illustrates.

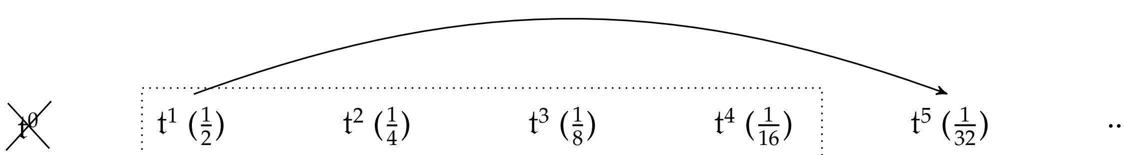

Figure 3. Flipping for Heads, after learning that the first flip landed tails. New convention — worlds which are crossed out are those that are incompatible with one's evidence.

This is a good result. Indeed—on the face of it, at least—this approach seems to vindicate No Gambler's Fallacy. 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 would now expect ANTICIPATION failures when modelling **Flipping for Both**. Surprisingly, this does not happen.

To model **Flipping for Both**, we'll first need to adjust our set of possibilities. Since the coin flipping procedure no longer terminates on the first flip landing on heads, we need to remove $\mathfrak{t}^0$. In its stead, we'll introduce worlds of the form $\mathfrak{h}^n$ — worlds in which the coin produces n heads in a row, followed by a tails. Assuming, again, that the relevant probabilities conform to the objective chances, and that one world is sufficiently more probable than another just in case it is at least 16 times more likely, we get the following diagram of **Flipping for Both** before the coin has been flipped:

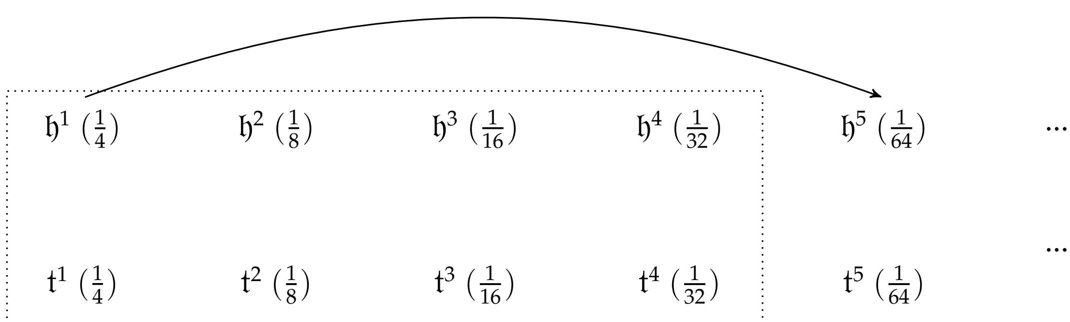

Figure 4. Flipping for Both, before any flips.

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. In sum, your strongest justified belief is that there will be at most a streak of 4 heads or 4 tails in a row.[23]

Matters get interesting on considering what you should believe upon learning that the first flip has landed tails (symmetric considerations apply to the

[23]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. Since we are not understanding these orderings of normality pre-theoretically, we could avoid this prediction by holding that, when modelling **Flipping for Both**, we just need to change our threshold for when one world is sufficiently more probable than another.

case in which you learn that the first flip has landed 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 on the current model. As the first flip has landed tails, all worlds in which the first flip landed heads are now inconsistent with your evidence. Accordingly, t^1 increases in probability to $\frac{1}{2}$ (as it now obtains if the next flip lands heads), 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:

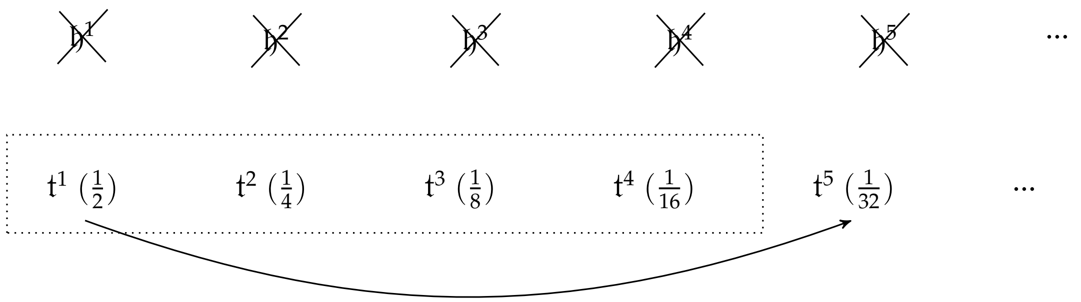

Figure 5. Flipping for Both, after first flip.

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 violate No Gambler's Fallacy. But this approach fails to extend this desirable feature to **Flipping for Both**. Here, the model predicts the *counterexamples* to No Gambler's Fallacy: on seeing the first flip land tails, you should now expect the first heads to occur sooner! So, insofar as we find this approach attractive because it seemed to track No Gambler's Fallacy, we now see that their approach fails to do so once applied to **Flipping for Both**. We therefore need an alternative approach.

3.2 Objective Normality — A Simple Model

It's instructive to diagnose why Lin and Kelly's approach fails to predict Anticipation failure in **Flipping for Both**.[24] As I see it, the key issue is that, on Lin and Kelly's approach, whether a world is among the most normal depends depends on how its probability compares to the probability of other worlds. That is, they understand normality as a *relative* notion.

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 if we reject this idea that normality is a relational matter; that is, the idea that how normal a world is depends on how its probability compares to the probability of other worlds. Alternatively, we can treat normality as an *objective* matter, judging the normality of a world by how its probability compares to some independent, fixed value. Doing so may allow for 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.

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

> **Objective Normality**
> A world w counts as among the most normal for agent S just in case, given S's evidence, the probability of w is at least τ ($0 \leqslant \tau \leqslant 1$).

Note that, while I have motivated Objective Normality through diagnosing Lin and Kelly's approach, the picture is provides of justified belief is also independently natural. In essence, Objective Normality tells us that one is

[24] As I outline in (Pearson ms, §4.1), the same diagnosis applies as to why Goodman and Salow's (2023; forthcoming) theory also fails to predict these failures of Anticiaption. The same diagnosis also applies to the theories defended in (Goldstein and Hawthorne 2021) and (Hong 2023).

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 one is 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, as endorsed by, for example, (Harsanyi 1985), (Lance 1995), (Lin 2013) and (Ross and Schroeder 2014).[25]

We can construct an attractive toy model of **Flipping for Both** in line with Objective Normality that can successfully predict the failure of Anticipation argued for in §2.2. (I'll then develop the approach in §3.3) This model can be effectively illustrating using bar charts, in the following manner. Each bar along the x-axis will represent a different possibility **Flipping for Both**. The height of each bar will represent that possibilities probability, measured along the y-axis. τ will then be 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 sufficiently normal. One is then justified in believing p just in case it is true in all of the shaded worlds. Setting $\tau = \frac{1}{16}$ (for the sake of a simple diagram), this is how things look in **Flipping for Both** before the coin has been flipped:

[25] 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 are important differences. Levi motivates his theory through epistemic decision theory. Due to this, Levi's theory predicts 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 — ends up making very different predictions to the theory I endorse here and cannot, for instance, make the desired predictions concerning **Flipping for Both**.

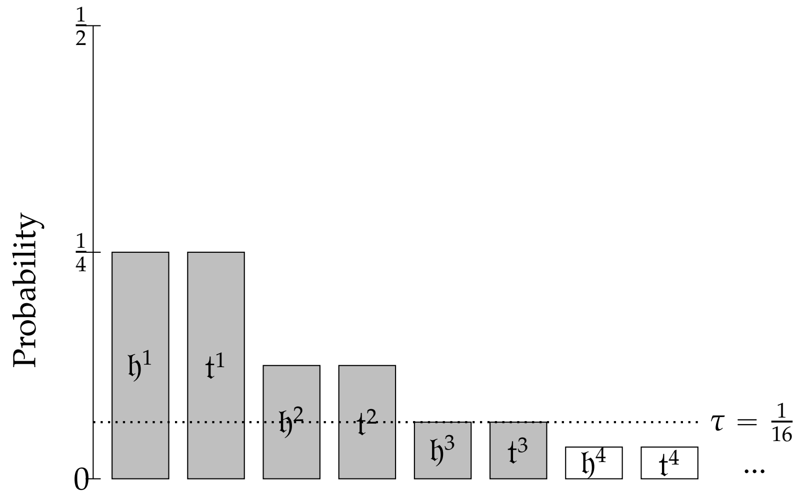

Figure 6. New model for **Flipping for Both** before any flips. Conventions — the height of each bar represents the probability of the world written inside the bar. Bars shaded in represent all and only those worlds consistent with what the relevant agent is justified in believing.

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 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:

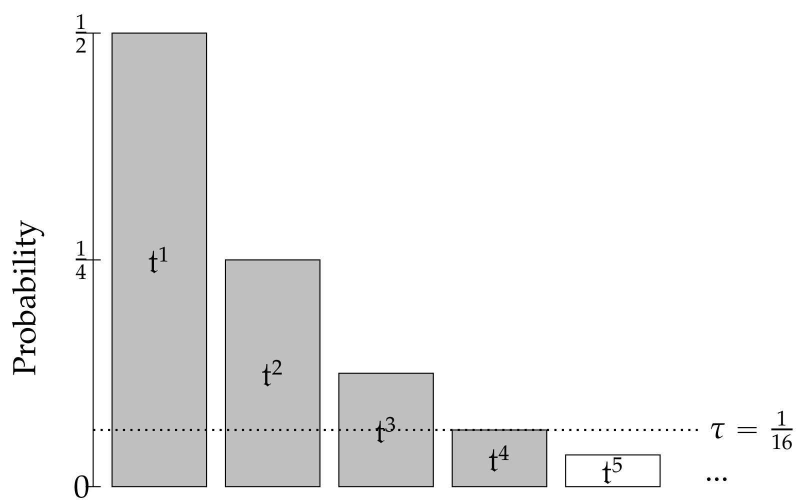

Figure 7. Alternative **Flipping for Both** model after first flip lands tails.

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. In turn, unlike with Lin and Kelly's approach, we predict no counterexamples to No GAMBLER'S FALLACY.

3.3 Objective Normality — Developing the View

Despite these attractive results, the simple model faces two issues that prevent it from forming the basis of a plausible theory of belief revision. I'll discuss these issues now and develop a more sophisticated version of the view that solves them.[26]

Issue 1: Justified Beliefs in Unlikely Propositions

The first issue is that the approach so far allows for an agent to be justified in believing p even if p is extremely unlikely. Consider, for instance, 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. Figure 8 illustrates.

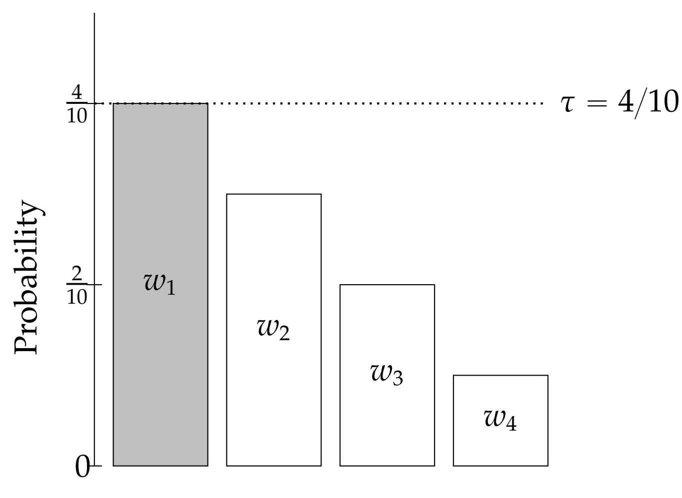

Figure 8. Illustration of issue 1.

[26]My presentation of these issues and my responses to them are slightly informal; (Pearson ms) provides the formal details.

A particularly acute version of this problem arises if *none* of the possibilities have a probability of at least τ. Suppose, for instance, that in Figure 8 τ is instead set at slightly above $\frac{4}{10}$. 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.[27]

How to solve this? I take the following idea to be quite natural. If none of the possibilities being considered by an agent have a probability of at least τ, rather than absurdly ruling out all those possibilities, the agent should instead concede that things are not as normal as one usually has a right to suppose. Nevertheless, the agent could still plausibly be justified in believing that they are in a world that is *pretty* normal, or perhaps *somewhat* normal, even if they are not justified in believing that things are *very* normal, so long as it is sufficiently likely that they are in a pretty normal/somewhat normal world.

The picture of justified belief this reply suggests can be intuitively expressed in procedural terms.[28] Roughly, the relevant agent should first examine all of the most normal worlds. If it is sufficiently likely that one of those worlds obtains, the agent can then justifiably believe they are in one of those worlds. If not, the agent should then examine all of the worlds that are at least *pretty* normal. If it is sufficiently like that one of these worlds obtains, the agent can then justifiably believe they are in a world that is at least pretty normal. If not, the agent should then examine all of the worlds that are at least *somewhat* normal, and so on, until the agent has found the highest degree of normal worlds such that it is sufficiently likely one of those worlds obtains.

We can implement this idea more precisely as follows. First, we introduce multiple thresholds τ_n, rather than a single threshold τ, where each threshold determines a different degree of normality. Let τ_1, for instance, represent the threshold of probability required for a world to possesses the highest degree of normality. τ_2 can then represent the threshold required for the next degree of normality, and so on. We'll call the set of worlds with a probability of at least τ_n the 'τ_n-normal' worlds. We'll then introduce another threshold, T, which any proposition must reach in order to be justifiably believed. Finally, we'll adjust our definition of justified belief so that, an agent is justified in believing they

[27]While theories of weak belief, such as those in (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.

[28](Hong 2023) uses a slightly different approach to model this same basic idea; however, his approach is incompatible with the failures of Anticipation I have argued for, unlike the approach I sketch just below.

are in a τ_n-normal world iff it is at least T-likely that they are in a τ_n-normal world.

For example, let's return to the case depicted in Figure 8. Instead of just having one threshold τ, set at $\frac{4}{10}$, we can instead suppose that while τ_1 is set at $\frac{4}{10}$, we also have τ_2 set at $\frac{3}{10}$ and τ_3 set at $\frac{2}{10}$. Setting T to $\frac{9}{10}$, we can see that the relevant agent is not justified in beleiving they are in a τ_1-normal world, as this is only $\frac{4}{10}$ths likely. Neither are they justified in believing they are in a τ_2-normal world, as this includes only w_1 and w_2, whose combined probability is only $\frac{7}{10}$. Instead, the strongest justified belief is that our agent is in a τ_3-normal world, which contains worlds w_1, w_2, and w_3, which have a joint probability of $\frac{9}{10}$. Figure 9 illustrates.

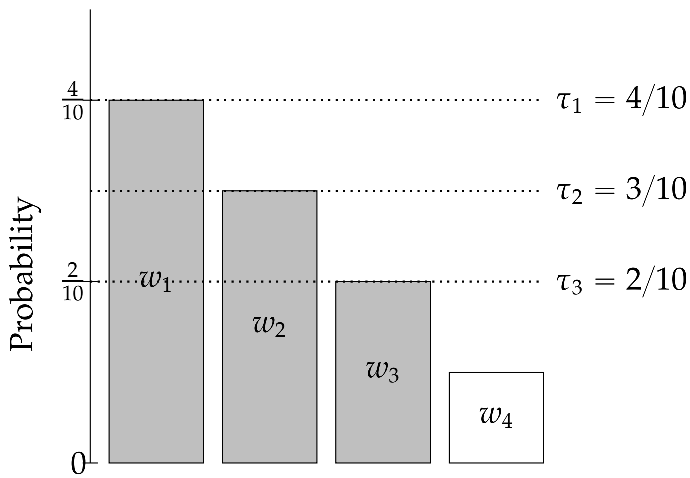

Figure 9. Illustration of solution to issue 1.

This approach not only solves the issue under discussion, it is moreover consistent with the predictions outlined in §3.2 regarding **Flipping for Both**, depicted in Figures 6 and 7 — just set, for instance, τ_1 to $\frac{1}{16}$ and T to, at most, $\frac{15}{16}$.

Issue 2: Influence of Irrelevant Information

For the second issue, consider:[29]

> **Dime or Nickel.** I am about to perform the coin-flipping procedure described in **Flipping for Both**. However, I am unsure whether to use a dime or a nickel for my coin. I decide to roll a 6-sided die to decide: I'll use a dime if it lands even and a nickel otherwise.

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

Plausibly, the matter of 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**. The problem is that the above approach 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 now be two worlds in which the coin lands tails 4 times, and each will have probability $\frac{1}{32}$. Assuming the threshold τ is at $\frac{1}{16}$, as it is in figure 6, this will mean you are permitted to have stronger beliefs in **Dime or Nickel** than you are in **Flipping for Both**: in the former, but not the latter, you can rule out the coin landing tails 4 times in a row. This is absurd.

The most promising response to problems of this kind is to follow various other authors in this literature, such as Lin and Kelly (2012), Leitgeb (2017),[30] Hong (2023) and Goodman and Salow (forthcoming) by endorsing a question-sensitive account of justified belief.[31] 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.[32] 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 possibilities.[33]

---

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

[31] See also (Holguín 2022), (Blumberg and Lederman 2020) and (Yalcin 2018) for other question-sensitive approaches to belief.

[32] 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)).

[33] Introducing question-sensitivity opens up another potential reply to my arguments against Anticipation in §2.2: perhaps I, objectionably, shift between questions such as *How many consecutive tails will there be?* and *How long will the opening consecutive sequence, of either heads or tails, be?* I discuss this reply in detail in the longer version of the paper (Pearson ms, §3.2.3). In short, I argue that all of my arguments still go through so long as we hold fixed the

Taking into account my responses to both issues, our developed theory of justified belief is therefore as follows. First, we have a set of possibilities, which is determined by the salient question, as outlined just above. Setting values for multiple thresholds τ_n, these possibilities are then ranked in terms of their probability. In general, all of the worlds with a probability of at least τ_n belong to the τ_n-normal worlds. What is one justified in believing? Let k be the smallest number such that the set of τ_k-normal worlds has a probability of at least T. One is justified in believing p iff p is true in all of the τ_k-normal worlds.

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 one can rule out sufficiently improbable possibilities.

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.

most fine-grained relevant question: *What sequence will the coin produce?*

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 is one tempting idea. Notice that they theory I defended in §3, as well as the theories I criticised — like Lin and Kelly's — encode a distinction between propositions you are merely justified in believing and propositions that are also a part of your *evidence*. Perhaps the above popular ideas that were assigned to justified beliefs in general should instead be restricted so as to only apply to those justified beliefs which are also part of your evidence. So, even if *belief* revision is not importantly connected to indicative conditionals, perhaps *evidence* revision is. Similarly, even if it is not in general permissible to use the propositions you are justified in believing as premises in reasoning, perhaps you can always use those propositions that are a part of your evidence as premises in reasoning.

This tempting idea comes at the cost of giving up on the importance of justified beliefs as such. If it is evidence, not justified beliefs generally, that play these important roles, then it appears to follow that justified beliefs less philosophically important than usually supposed, if the are philosophically important at all.[34] I'm inclined to resist this conclusion, but I cannot deny that my arguments forge a new path to reaching it.

Can we, instead, learn to live with the awkward consequences of accepting these ideas about belief alongside failures of ANTICIPATION? Here's one promising avenue. We can connect my arguments against ANTICIPATION to cases of "iteration" or "introspection" failures of belief. For our purposes, these are cases in which an agent has justification to believe a proposition p, without having justification to believe they have justification to believe p. Interestingly, **Flipping for Both** is arguably a case where this happens. Suppose that in **Flipping for Both** you must leave open a streak of 19 heads/tails in a row, but

[34]Recall, for example, that a cornerstone of Williamson's (2000, ch. 9) influential knowledge-first approach was the thesis that one's evidence is all and only the propositions one knows. Williamson recognises that if there are some propositions one knows that are not part of one's evidence, there is at least some sense in which it is evidence, not knowledge, that comes first.

you are justified in believing there won't be a streak of 20. Even if that's so, it will still be extremely difficult for you to distinguish your actual case from a case in which you must instead leave open a streak 20 heads/tails in a row and are only justified in believing there won't be a streak of 21. Accordingly, your justified belief that there won't be a streak of 20 heads/tails in a row—the proposition for which ANTICIPATION fails—will be such that you do not have sufficient justification to believe you are justified in having that belief.

If this is right, we can offer the following explanation of what is objectionable about agents that make the outlined bizarre assertions or decline free evidence, that follows a picture endorsed by (Williamson forthcoming) and (Carter and Hawthorne forthcoming). Though these agents are acting in accordance with their justified beliefs, and thus in a way that is epistemically *permitted*, they are nonetheless acting is ways that are epistemically *risky*. That is, from their perspective, they are not in a position to justifiably believe they are acting in a way that is epistemically permitted, since they cannot justifiably believe they are acting in accordance with their justified beliefs. This approach is of course in need for further elaboration. But if it works, then perhaps we can learn to live without ANTICIPATION.[35]

References

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.

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

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

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

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.

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 (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*.

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.

Huber, Franz (2013). "Belief Revision I: The Agm Theory". In: *Philosophy Compass* 8.7, pp. 604–612. DOI: 10.1111/phc3.12048.

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.

Kripke, Saul A. (2011). "On Two Paradoxes of Knowledge". In: *Philosophical Troubles. Collected Papers Vol I*. Ed. by Saul A. Kripke. Oxford University Press.

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.

Lewis, David (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.

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.

— (2019). "Belief Revision Theory". In: *The Open Handbook of Formal Epistemology*. Ed. by Richard Pettigrew and Jonathan Weisberg. PhilPapers Foundation, pp. 349–396.

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.

Pearson, Joshua Edward (ms). "Belief Revision Revised". Manuscript - joshuaedwardpearson.com.

Ramsey, Frank (1926). "Truth and Probability". In: *Philosophy of Probability: Contemporary Readings*. Ed. by Antony Eagle. Routledge, pp. 52–94.

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 (forthcoming). "Fallibility and Dogmatism". In: *Australasian Journal of Philosophy*. DOI: 10.1080/00048402.2024.2381692.

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.

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.

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.