



Coin trials

Martin Smith

Department of Philosophy, University of Edinburgh, Edinburgh, UK

ABSTRACT

According to the JUSTIFIED FAIR COINS principle, if I know that a coin is fair, and I lack justification for believing that it won't be flipped, then I lack justification for believing that it won't land tails. What this principle says, in effect, is that the only way to have justification for believing that a fair coin won't land tails, is by having justification for believing that it won't be flipped at all. Although this seems a plausible and innocuous principle, in a recent paper Dorr, Goodman and Hawthorne use it in devising an intriguing puzzle which places all justified beliefs about the future in jeopardy. They point out, further, that one very widespread theory of justification predicts that JUSTIFIED FAIR COINS is false, giving us additional reason to reject it. In this paper, I will attempt to turn this dialectic around. I will argue that JUSTIFIED FAIR COINS does not inevitably lead to scepticism about the future, and the fact that it is incompatible with a widespread theory of justification may give us reason to doubt *the theory*. I will outline an alternative theory of justification that predicts both that JUSTIFIED FAIR COINS is true and that we have justification for believing much about the future.

ARTICLE HISTORY Received 18 November 2016; Accepted 13 September 2017

KEYWORDS Probability; normalcy; justification; fair coin; sorites paradox

1. A puzzle about justification

Consider the following principle:

If I know that a coin is fair and, for all I know, it is going to be flipped then, for all I know, it is going to land tails.

Dorr, Goodman and Hawthorne (2014) call this principle FAIR COINS. What this principle says, in effect, is that the only way to know that a fair coin won't land tails is to know that it won't be flipped at all. If we put aside possibilities such as clairvoyance, time travel and 'cheesy' modes of presentation¹ then FAIR COINS seems a very appealing principle. Now suppose there are 1000 fair coins arrayed

CONTACT Martin Smith  Martin.Smith@ed.ac.uk

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

before me which I'm about to start flipping. If the first coin lands tails, I'll flip the second. If the second coin also lands tails, I'll flip the third and so on. As soon as a coin lands heads, I stop flipping and the remaining coins are destroyed. Suppose that, in actual fact, the second coin is going to land heads and that will be the end of it – but imagine that I'm just about to start the process. If I know everything about the set-up, what can I know, at this point, about the outcome?

Before I start flipping, it is extremely objectively likely that coin 1000 will not be flipped. After all, this would require throwing 999 tails in a row, and the objective probability of this is 0.5^{999} – (far) less than 1 divided by the number of atoms in the known universe.² But can I *know* that coin 1000 won't be flipped? Given my knowledge of the set-up, if I know that coin 1000 won't be flipped, then I know that coin 999 won't land tails. Given FAIR COINS, it follows that I know that coin 999 won't be flipped. But then, given my knowledge of the set up, I know that coin 998 won't land tails and so on. After 1000 iterations we derive the absurdity that I know that coin 1 won't be flipped. If FAIR COINS holds then I cannot know that coin 1000 won't be flipped, even though this is true and extremely objectively likely.

As Dorr, Goodman and Hawthorne point out, a similar principle to FAIR COINS can be used to derive the result that I can't even have *justification for believing* that coin 1000 won't be flipped.

If I know that a coin is fair and I lack justification for believing that it isn't going to be flipped, then I lack justification for believing that it isn't going to land tails.

Dorr, Goodman and Hawthorne call this principle JUSTIFIED FAIR COINS.³ What this principle says, in effect, is that the only way to have justification for believing that a fair coin won't land tails is to have justification for believing that it won't be flipped. Once again, if we set aside clairvoyance, time travel, cheesy modes of presentation etc., JUSTIFIED FAIR COINS seems a very appealing principle. Using this principle, Dorr, Goodman and Hawthorne's argument can be run as before, but with justification in place of knowledge: Given what I know about the set-up, if I have justification for believing that coin 1000 won't be flipped, then I have justification for believing that coin 999 won't land tails. Given JUSTIFIED FAIR COINS, it follows that I have justification for believing that coin 999 won't be flipped. But then, given my knowledge of the set up, I have justification for believing that coin 998 won't land tails and so on. After 1000 iterations we derive the absurdity that I have justification for believing that coin 1 won't be flipped. If JUSTIFIED FAIR COINS holds, then I cannot have justification for believing that coin 1000 won't be flipped, even though it is true and extremely objectively likely.

Although JUSTIFIED FAIR COINS seems an appealing principle, surely this puzzle forces us to give it up – surely we can't deny that I have justification for believing that coin 1000 won't be flipped. And, even if we can bring ourselves to deny this, things may get even worse. Following Dorr, Goodman and Hawthorne (Section 3), suppose I'm looking at a particular leaf on a particular maple tree at the beginning of autumn. I know that the tree will soon start to shed its leaves

and by the end of winter is almost certain to be bare. Consider the proposition that, by the end of winter, the leaf I'm looking at will have fallen from the tree. This is not a proposition about an artificial or esoteric set up but is, rather, the kind of thing we could easily imagine ourselves believing about the future.

But the period between now and the end of winter can be divided up into hour-long intervals and I know that, for any given hour, if the leaf is still on the tree at the beginning of the hour then there is a high objective probability that it will still be on the tree at the end of the hour. With this in mind, consider the following principle, which we might call JUSTIFIED AUTUMN LEAF:

If I know that leaf shedding is a chancy process then, for any number n , if I lack justification for believing that the leaf will have fallen by the beginning of the n th hour, then I lack justification for believing that the leaf will have fallen by the end of the n th hour.

If I have justification for believing that the leaf will have fallen by the end of winter then, by this principle, I must have justification for believing that the leaf will have fallen by an hour before the end of winter and, by applying the principle again, I must have justification for believing that the leaf will have fallen by an hour before that and so on. By applying the principle enough times we can derive the absurdity that I have justification for believing that the leaf will fall from the tree in the next instant.

If we accept JUSTIFIED FAIR COINS, we are under some pressure to accept JUSTIFIED AUTUMN LEAF and, if we do, we are forced to deny that I have justification for believing that the leaf will have fallen by the end of winter.⁴ But if we deny this, then we really do seem to be inviting a more pervasive scepticism about the future – a view on which most, if not all, of the things we believe about the future are unjustified. Furthermore, we could hardly prevent such scepticism from infecting some beliefs about the present and past as well. It seems, then, that we have a powerful reason for rejecting JUSTIFIED FAIR COINS.

In addition, as Dorr, Goodman and Hawthorne point out (n8), there is a widespread conception of epistemic justification that provides further, independent reason for rejecting JUSTIFIED FAIR COINS. On this conception, one has justification for believing a proposition P just in case, given one's evidence, it would be *unlikely* for P to be false. Elsewhere I have termed this the *risk minimisation* theory of justification – it makes justification a matter of minimising the risk of error. To make this idea a little more formal, suppose we set a threshold t greater than, but close to, 0 to represent the maximum tolerable risk of error. According to the risk minimisation theory, I have justification for believing a proposition P just in case the probability that P is false, given my evidence, is less than t .

If the risk minimisation theory is correct, then JUSTIFIED FAIR COINS is false – the risk of falsity to which I expose myself in believing that a fair coin won't land tails is lower than the risk of falsity to which I expose myself in believing that it won't be flipped. By plugging the risk minimisation theory into JUSTIFIED FAIR COINS we derive the following principle:

If I know that a coin is fair and the probability, given my evidence, that it will be flipped is greater than t , then the probability, given my evidence, that it will land tails is greater than t .

It is obvious that this principle is false – the probability of a fair coin in the sequence landing tails is lower than the probability of it being flipped and, as such, the two values could easily sit either side of a numerical threshold.

Suppose we fix the threshold t as equal to 0.1 – so I have justification for believing anything which, given my evidence, has less than a 0.1 probability of being false. In Dorr, Goodman and Hawthorne's scenario, the probability, given my evidence, that coin 2 will be flipped is 0.5, that coin 3 will be flipped is 0.25 and that coin 4 will be flipped is 0.125. The probability, given my evidence, that coin 2 will land tails is 0.25, that coin 3 will land tails is 0.125 and that coin 4 will land tails is 0.0625. The risk minimisation theory straightforwardly predicts, then, that I have justification for believing that coin 4 won't land tails, but I lack justification for believing that coin 4 won't be flipped.⁵

Here are the probabilities of each coin being flipped and of landing tails, respectively:

	COIN					
	1	2	3	4	5	6...
Flipped	1	0.5	0.25	0.125	0.0625	0.03125
Tails	0.5	0.25	0.125	0.0625	0.03125	0.015625

Evidently, the error threshold t at crossed at coin 4.

At this point, the case against JUSTIFIED FAIR COINS may seem decisive – even overdetermined. Not only does the principle appear to lead to scepticism about the future, it clashes with a widely held, independently motivated, theory of epistemic justification. And yet, as powerful as these considerations seem, there are still reasons for being somewhat uneasy. JUSTIFIED FAIR COINS *does* seem like an appealing principle, and, in a way, we are still no closer to understanding why this should be. Indeed, the fact that the risk minimisation theory converts this principle into such a patent falsehood is, perhaps, something that should give us pause. It is, of course, right and proper that a theory of justification should prompt us to revise certain pretheoretic judgments – but the risk minimisation theory makes it difficult to see why this principle would have ever seemed attractive in the first place.

Further, consider the following, which we might call JUSTIFIED BIASED COINS:

If I know that a coin is heavily biased in favour of tails and, I don't have justification for believing that it isn't going to be flipped then I don't have justification for believing that it isn't going to land tails.

This principle might seem even more appealing than JUSTIFIED FAIR COINS. If I know that a coin is very heavily biased towards tails, it would be very odd to believe that it won't land tails, while remaining agnostic about whether it will be flipped – and it's difficult to accept that such a combination of attitudes could ever be justified. But the risk minimisation theory will also translate *this*

into an obvious falsehood. Even if a coin is heavily biased in favour of tails, the probability of it landing tails will still be slightly lower than the probability of it being flipped and these two values could easily sit either side of a threshold.

2. Two theories of justification

In previous work (Smith 2010, 2016), I've argued against the risk minimisation theory of justification, and in favour of an alternative that I've termed the *normic* theory. According to the risk minimisation theory, one has justification for believing a proposition *P* just in case, given one's evidence, it would be unlikely for *P* to be false. According to the normic theory, one has justification for believing a proposition *P* just in case, given one's evidence, it would be *abnormal* for *P* to be false.

Sometimes when we describe a situation as 'abnormal' we mean only that it is relatively rare or infrequent. If this is how we are understanding the notion of normalcy, then the normic theory looks very similar to the risk minimisation theory, and is not obviously an alternative at all. Other times, though, when we describe a situation as abnormal, we are not making a claim about frequencies. Rather, we are flagging the situation as an exception to a pattern or regularity – as something that would need *special explanation* if it were to obtain. If the lights in my house suddenly start to flicker, or my car fails to start when I turn my key in the ignition and I remark 'that's not normal', I'm not simply pointing out that this is something rare or infrequent – what I'm saying is that there needs to be some explanation for what is occurring. This is the understanding of normalcy that I have in mind in putting forward the normic theory – an understanding that will make it into something recognisably different from the risk minimisation theory.

Suppose I'm wondering whether the milk has gone off. I sniff it and I can't smell anything unusual, and I check the use by date and see that it's still a few days away. Clearly this evidence makes it very *unlikely* that the milk is off – but this is not its only effect. If the milk were off, in spite of the evidence, then some special explanation would be needed – I've lost my sense of smell or the odour has been masked and the wrong date was stamped on the milk or the milk has not been properly refrigerated etc. Given my evidence it would be unlikely *and* abnormal for the milk to be off. Both the risk minimisation and normic theories predict, in this case, that I have justification for believing that the milk is not off – but they offer different diagnoses for why this is.

While the risk minimisation and normic theories will often agree in their predictions, they do not always do so. The normic theory seems, for instance, to offer a very different perspective on JUSTIFIED FAIR COINS. By plugging the normic theory into this principle we derive the following:

If I know that a coin is fair and it is normal, given my evidence, for it to be flipped then it is normal, given my evidence, for it to land tails.

Call this NORMIC FAIR COINS. Given the intended reading of 'normal' this principle seems clearly *true*. If I know that a coin is fair and the situation in which it is flipped requires no special explanation, given my evidence, then the situation in which it lands tails requires no special explanation, given my evidence.

It may even be possible to motivate NORMIC FAIR COINS without appealing to any substantial theory of normalcy. Given that a coin is fair, it would be equally normal for it to be flipped and land heads and for it to be flipped and land tails. Arguably, this is part of the analysis of what it means for a coin to be 'fair' – if it is more normal for a coin to be flipped and land heads than for it to be flipped and land tails or vice versa, then it is not a fair coin. But, given a few principles about the formal properties of normalcy, NORMIC FAIR COINS can be 'proved' from this trivial starting point.⁶ In the Appendix I describe a simple formal model of normalcy in which this derivation can be set out – but I leave the discussion informal here.

Since the normic theory predicts that JUSTIFIED FAIR COINS is true, it will play into Dorr, Goodman and Hawthorne's hands and predict that, in the scenario they describe, I lack justification for believing that coin 1000 won't be flipped. But this result should come as no particular surprise – given my evidence, it may be extremely unlikely for me to throw 999 tails in a row, but this outcome is not abnormal in the sense of requiring special explanation.⁷ Put aside Dorr, Goodman and Hawthorne's particular set-up for a moment and imagine a simpler case in which I'm about to throw 999 fair coins and note the result. Of the 2^{999} possible sequences that I could get, it's clear that none of these is abnormal, or requires any special explanation. The sequences are all on a par. When it comes to a 'patterned' sequence like 999 tails, we might be inclined to *think* that some special explanation would be needed for this – but, in reality, this sequence could come up just as easily as any other (Smith *forthcoming*).⁸ In a case like this, the normic theory predicts that I should be open minded about how the coins will land – I don't have justification for ruling out any outcomes in advance.

In Dorr, Goodman and Hawthorne's set-up, in which the sequences terminate with the first head, there are far fewer possible outcomes and the sequence of 999 tails is the least likely of them. But this sequence is no less normal, and requires no more explanation, than it does in the simpler case.⁹ While I may be justified in believing that this outcome is extremely unlikely to obtain, according to the normic theory, I lack justification for believing that it won't obtain.

What then of Dorr, Goodman and Hawthorne's claim that this verdict threatens a slide into scepticism about the future? Does the normic theory carry such a threat? What does it predict, for instance, about the autumn leaf example? I will build up to answering this question by considering two related cases: Suppose I've set up my computer to sound an alarm at some randomly determined moment within the next two hours. Presumably, I'm justified in believing that my computer will sound an alarm some time during the next two hours.

This case has a number of structural features in common with the autumn leaf example. After all, the next two hours can be divided up into, say, minute-long intervals and I know that, for any given minute, if the alarm has not sounded by the beginning of the minute, then there is a high objective chance that it will not have sounded by the end of the minute. Consider, then, the following principle:

JUSTIFIED ALARM

If I know that the alarm is determined by a chancy process then, for any number n , if I lack justification for believing that the alarm will have sounded by the beginning of the n th minute, then I lack justification for believing that the alarm will have sounded by the end of the n th minute.

While this principle seems very solid for the first minute, the second minute, the third minute ... the 120th minute is clearly special. It's consistent with my computer functioning properly that the alarm has not sounded by the beginning of the 120th minute. It is not consistent with my computer functioning properly that the alarm has not sounded by the end of the 120th minute. It's plausible at least that this minute serves as a counterexample to JUSTIFIED ALARM – I lack justification for believing that the alarm will have sounded by the beginning of the 120th minute, but I have justification for believing that it will have sounded by the end of the 120th minute.

This impression is further cemented by considering the normic translation of JUSTIFIED ALARM:

NORMIC ALARM

If I know that the alarm is determined by a chancy process then, for any number n , if it is normal, given my evidence, for the alarm not to have sounded by the beginning of the n th minute, then it is normal, given my evidence, for the alarm not to have sounded by the end of the n th minute.

Given the way I set up my computer, the instant at which we reach the end of the 120th minute is the instant at which a failure of the alarm to sound becomes abnormal and requires explanation of some kind – the computer has lost power, someone has taken to it with a hammer etc.

What this example shows, at the very least, is that the normic theory will not lead to an all-encompassing scepticism about the future. My justification for believing that the alarm will sound within the next two hours seems to be safe from a Dorr, Goodman and Hawthorne-style argument. One might complain, though, that this case is rather artificial and contrived – it's rare to have such precise information about the time period during which an event might occur. Here is a more realistic, but closely related, case: Suppose my roommate calls and tells me that he'll be home within the next two hours. Consider the following principle:

NORMIC ARRIVAL TIME

If I know that arriving home is a chancy process then, for any number n , if it is normal, given my evidence, for my roommate to have not arrived home by the

beginning of the n th minute, then it is normal, given my evidence, for my roommate to have not arrived home by the end of the n th minute.

NORMIC ARRIVAL TIME looks a lot more secure than NORMIC ALARM. It doesn't seem right that, at the end of the 120th minute, my roommate's failure to arrive home would suddenly become abnormal and require explanation. In fact, it's strange to think that there's *any* minute during which this happens.

Unlike NORMIC ALARM, we can't find a clear counterexample to NORMIC ARRIVAL TIME – but this is not to say that we should accept it. If my roommate has not arrived within 20 min or within 1 h, then there may be nothing abnormal about that. But if my roommate has still not arrived within 10 or 24 h then that clearly *would* be abnormal and some explanation would be needed. NORMIC ARRIVAL TIME is, however, inconsistent with these judgments. If the situation in which my roommate has not arrived by the end of the first minute is not abnormal then, by repeatedly applying NORMIC ARRIVAL TIME, it follows that the situation in which my roommate has not arrived within 24 h is not abnormal. This, of course, is nothing other than an instance of the *sorites paradox*.

In the classic presentation of the paradox, we begin with a base case – a man with 0 hairs on his head is bald (it is normal for my roommate to have not arrived home by the end of minute 1/beginning of minute 2). We then introduce a soritical principle – for any number n , if a man with n hairs on his head is bald, then a man with $n + 1$ hairs on his head is bald (if it is normal for my roommate to have not arrived home by the beginning of the n th minute, it is normal for my roommate to have not arrived home by the end of the n th minute/beginning of the $n + 1$ th minute). From these two premises, an absurd conclusion follows – a man with 1,000,000,000 hairs on his head is bald (it's normal for my roommate to have not arrived home within 24 h).¹⁰

At first, NORMIC ARRIVAL TIME might look more like NORMIC FAIR COINS than NORMIC ALARM. After all, there is a clear counterexample to NORMIC ALARM, while NORMIC ARRIVAL TIME and NORMIC FAIR COINS seem to be counterexample-proof. On closer inspection, though, these two principles turn out to be very different. NORMIC FAIR COINS can be given a sound motivation – and perhaps even be motivated by an *analysis* of what it means for a coin to be 'fair'. NORMIC ARRIVAL TIME is not motivated by an analysis of anything – its appeal lies in that which makes all soritical principles seductive – our reluctance to impose a sharp boundary on vague expressions. Almost all contemporary treatments of the sorites paradox involve denying that soritical principles are perfectly true, in spite of their undeniable appeal, and lack of clear counterexamples.¹¹ I am inclined to follow this approach, but whatever view we take of soritical principles, the fact that they lead to trouble when combined with the normic theory is evidently no reflection on the theory – they lead to trouble on their own.

In the case described, the normic theory predicts that I have justification for believing that my roommate will have arrived home within 24 h – and our only reasons for thinking otherwise are disreputable soritical reasons. What this

example shows is that the normic theory will not lead to a widespread scepticism about the future. My justification for believing that my roommate will be home within 24 h – typical of much of what we have justification for believing about the future – is safe from a Dorr, Goodman and Hawthorne-style argument.

Return, finally, to the autumn leaf example. If we plug the normic theory into JUSTIFIED AUTUMN LEAF we obtain the following:

NORMIC AUTUMN LEAF

If I know that leaf shedding is a chancy process then, for any number n , if it's normal, given my evidence, for the leaf to have not fallen by the beginning of the n th hour, then it's normal, given my evidence, for the leaf to have not fallen by the end of the n th hour.

Here is one compelling thought: The leaf shedding process serves a function or purpose for the tree and, absent external interference, will unfold in a particular, predetermined way. If the leaf I'm looking at were still on the tree at the end of winter, then this would represent a breakdown in the process and, given my evidence, would require explanation in terms of some anomaly with the season, the immediate environment, the tree or the leaf. If we're struck by this thought then, given the normic theory, we should take ourselves to have justification for believing that the leaf will not be on the tree at the end of winter. On this way of thinking, we should regard NORMIC AUTUMN LEAF as nothing more than a soritical principle, with the associated deceptive allure.¹²

Here is an alternative thought: The leaf shedding process is essentially random, in the same sense as a series of coin flips, and each of the possible outcomes of the process, including those in which the leaf remains on the tree at the end of winter, is equally normal. If we're struck by this thought then, given the normic theory, we should not take ourselves to have justification for believing that the leaf won't be on the tree at the end of winter – the most we would have justification for believing is that it is highly likely that the leaf won't be on the tree at the end of winter. On this way of thinking, NORMIC AUTUMN LEAF will be trivially true.¹³

Each thought can seem appealing, in a certain frame of mind – but the crucial point, for present purposes, is just this: If the circumstance in which the leaf is still on the tree at the end of winter would be abnormal, given my evidence, then the normic theory predicts that I have justification for believing that the leaf will have fallen by the end of winter. If the circumstance in which the leaf is still on the tree at the end of winter would be normal, given my evidence, then the normic theory predicts that I lack justification for believing that this won't happen. But whatever the case, it was obvious *from the get-go* that the predictions of the normic theory would hinge on this very issue – the NORMIC AUTUMN LEAF principle is a wheel turning idly. That is, whatever sceptical liabilities there are in the normic theory, it wears them on its sleeve – there is no latent sceptical threat in the theory that a Dorr, Goodman and Hawthorne-style argument can draw out.

3. Conclusion

JUSTIFIED FAIR COINS is an appealing principle. And yet, Dorr, Goodman and Hawthorne argue that the principle is troublesome, ultimately leading us to scepticism about the future. The principle is also invalidated by a widespread and well-motivated theory of justification – the risk minimisation theory. This observation might be thought to clinch the case against JUSTIFIED FAIR COINS and to be a point in favour of the risk minimisation theory. In this paper I have attempted to subvert this picture of the dialectical situation. I have argued that JUSTIFIED FAIR COINS does not inevitably lead to scepticism about the future. I have outlined a theory of justification – the normic theory – which turns out to validate JUSTIFIED FAIR COINS, and on which scepticism about the future is not inevitable. In light of this, the fact that the risk minimisation theory invalidates JUSTIFIED FAIR COINS, and does so in a way that seems to make its appeal inexplicable, might even be seen as a reason to question the theory.

Some of what I say here has implications for the original FAIR COINS principle on which Dorr, Goodman and Hawthorne focus most of their attention. In the scenario that Dorr, Goodman and Hawthorne envisage, FAIR COINS will lead us to the result that I cannot know that coin 1000 won't be flipped, in spite of the fact that this is true and extremely objectively and evidentially likely. From the perspective of the normic theory, though, this provides no decisive reason to reject the principle. The idea that we are unable to know certain highly likely truths about the future is one that a normic theorist should already be quite accustomed to. After all, some highly likely truths about the future are such that, given our evidence, their falsity would involve no departure from normal circumstances. These will be truths that we cannot justifiably believe and, assuming that knowledge requires justification, cannot know either. The truth that coin 1000 won't be flipped is one such.

Notes

1. Some coins have the property of being flipped at some point during their existence, but never landing tails. Suppose I declare that 'Headsy' is the name of the next coin to be minted that has this property. Having introduced the name in this way, it looks as though I'm in a position to know that Headsy will be flipped but won't land tails (see Dorr, Goodman and Hawthorne 2014, 277–278, for related discussion see Hawthorne and Lasonen-Aarnio 2009, Section 3).
2. I'm assuming here that a fair coin has a 0.5 objective chance of landing tails and a 0.5 objective chance of landing heads when flipped. Realistically, there are other outcomes (the coin landing on its side, the coin failing to land) that have non-zero chances – but I follow Dorr, Goodman and Hawthorne (and the probability textbooks) in abstracting away from this.
3. The principle that Dorr, Goodman and Hawthorne consider is slightly different to the one I have written here. Their principle is: If I *have justification for believing* that a coin is fair and I don't have justification for believing that it isn't going to be flipped then I don't have justification for believing that it isn't going to land tails.

Assuming that justification is necessary for knowledge, this is logically stronger than the principle that I have termed JUSTIFIED FAIR COINS. The weaker principle is in fact sufficient to drive the sceptical problem to which Dorr, Goodman and Hawthorne draw attention. It is this principle on which I will focus.

4. Both JUSTIFIED FAIR COINS and JUSTIFIED AUTUMN LEAF could be derived from the following principle, the justification-equivalent of the principle that Dorr, Goodman and Hawthorne term POSSIBLE FUTURE UNLIKELIHOOD:

If I lack justification for believing that *P* doesn't and won't have a substantial objective probability of being false, then I lack justification for believing *P*.

Call this JUSTIFIED POSSIBLE FUTURE UNLIKELIHOOD. We might suppose that, when a fair coin is flipped, there is a substantial objective probability that it will land tails. If I lack justification for believing that a coin won't be flipped, then I lack justification for believing that there won't be a substantial objective probability of it landing tails in which case, using JUSTIFIED POSSIBLE FUTURE UNLIKELIHOOD, we can derive the result that I lack justification for believing that the coin won't land tails.

Similarly, we might suppose that, if a leaf has not fallen from the tree by the beginning of the *n*th hour there is a substantial objective probability that it will not have fallen from the tree by the end of the *n*th hour. If I lack justification for believing that the leaf will have fallen from the tree by the beginning of the *n*th hour then I lack justification for believing that there won't be a substantial objective probability that the leaf will not have fallen from the tree by the end of the *n*th hour in which case, using JUSTIFIED POSSIBLE FUTURE UNLIKELIHOOD we can derive the result that I lack justification for believing that the leaf will have fallen from the tree by the end of the *n*th hour. One might suggest that JUSTIFIED FAIR COINS derives its plausibility from the more general JUSTIFIED POSSIBLE FUTURE UNLIKELIHOOD in which case JUSTIFIED FAIR COINS and JUSTIFIED AUTUMN LEAF will effectively come as a package (Dorr, Goodman and Hawthorne, Section 5, speculate that the knowledge-equivalents of these principles may stand in some such relation). I will return to this in n12.

5. Dorr, Goodman and Hawthorne focus, in their discussion, on *objective*, rather than evidential, probabilities. I am supposing here that, at the point before any coins are flipped, the objective and evidential probabilities of the various possible outcomes coincide with each other, and with what a textbook probability calculation would suggest. But there are views that predict otherwise. If, for instance, we adopted the knowledge account of evidence (Williamson 2000, chap. 9) on which one's evidence equals the sum total of one's knowledge, then we couldn't assign evidential probabilities in this case without already taking a view on Dorr, Goodman and Hawthorne's *first* puzzle, involving the FAIR COINS principle. If we endorsed a non-concessive approach to that puzzle on which I do know that I won't flip 999 tails in a row then, according to the knowledge account, this will be part of my evidence, and thus will have an evidential probability of 1, as opposed to $1 - 0.5^{999}$.

While this might seem a rather small difference, it will have some very unusual downstream consequences. If it's certain that I won't throw 999 tails in a row, then there must be at least one coin in the sequence such that the evidential probability of it landing tails, given that all previous coins have landed tails, is 0 – that is, a coin that is *certain* to land heads, conditional on all previous coins landing tails. And this is the case even though it is also certain, given my evidence, that the coin in question is fair and that it is completely unaffected by any of the previous flips etc. (for a penetrating discussion of some related issues see Bacon, 2014). I think this idea is close to incoherent, and constitutes a powerful objection

to combining the knowledge account of evidence with a non-concessive solution to the knowledge-puzzle. In any case, I don't consider this combination of views in the main text, but in n7 I mention one way in which it would significantly bear upon the subsequent discussion.

6. The standard definition of a 'fair coin' makes no explicit mention of normalcy, of course, and would perhaps go something like this: A coin is fair just in case it has an equal objective probability of landing heads or tails when flipped. And yet, it does sound contradictory to say 'This coin is fair but, when we flip it, it would be more normal for it to land heads than tails', suggesting either that this is precluded by the standard definition (in virtue of certain necessary connections between normalcy and objective probability) or that the standard definition needs to be supplemented. I won't pursue this further here. Even if one cleaves to the standard definition, and denies any connections with normalcy, NORMIC FAIR COINS can still be motivated by appealing to the particular conception of normalcy developed in the main text.
7. As discussed in n5, if one endorsed a non-concessive solution to the initial knowledge-puzzle, and embraced the knowledge account of evidence, then the proposition that I won't throw 999 tails in a row will be a part of my evidence and, thus, the falsity of this proposition *would* be abnormal, given that evidence (impossibility being a limiting case of abnormality). I have already offered one objection to this combination of views and, while it's not my aim here to argue against it at length, I will note one further reservation: When it comes to Dorr, Goodman and Hawthorne's set-up, facts about the evidential probabilities of the various outcomes and, to a lesser extent, about the evidential normalcy of the various outcomes seem to be the *clearest* features of the case – amongst the few features that we have at our disposal in attempting to solve the puzzles that the case raises. By insisting that we can't even *access* these features until we have offered a brute response to the knowledge-puzzle, we preclude the possibility of any satisfying solution.
8. For contrasting views, on which the different sequences differ with respect to their normalcy, and on which a sequence of 999 tails would count as highly abnormal, see Dutant (2016) and Goodman and Salow (forthcoming). Dutant and Goodman and Salow are working with different conceptions of normalcy to the one that I have set out. In so far as these conceptions are regarded as viable, they will put pressure on the claim that NORMIC FAIR COINS can be extracted purely from the meaning of what it is for a coin to be 'fair', along with certain formal principles. I will return to these views in the appendix.
9. If we had not investigated or inspected the coins being used then, while a run of 999 tails would not *require* a special explanation, it would make it *very likely* that there is a special explanation – that the coins are double-tailed or weighted etc. In the case that Dorr, Goodman and Hawthorne describe, however, it is part of the background evidence that all the coins are *fair*. Holding this evidence fixed, I would be inclined to deny that a special explanation would even be likely in the event of 999 tails.
10. Classic presentations of the sorites paradox may encourage the impression that vague expressions represent an arbitrary cut-off on some underlying scale, that is then exploited in constructing a sorites series for the expression. What, then, is the underlying scale on which 'requires explanation' represents an arbitrary cut-off? In fact, not all vague, sorites-susceptible expressions fit this model. 'Knows', for instance, is a vague term, and one can certainly construct sorites series for it, but it doesn't represent an arbitrary cut-off on some underlying scale (or so the conventional view would have it). While the members of a sorites series for

'knows' or for 'requires explanation' must, of course, vary along some dimension, it needn't have any essential connection to the meanings of these terms. I won't pursue this further here.

11. On epistemicist treatments of the paradox there will be one instance of a soritical principle that is false, though we cannot know which instance this is. On supervaluationist treatments, there will be a range of instances of a soritical principle that are not 'super-true' – that is, true according to all permissible sharpenings of the vague expression involved – but not super-false either. On this view, though, the soritical principle itself *will* count as super-false – since on every permissible sharpening of the vague expression, there will be one false instance. On 'many valued' treatments, the soritical principle will, at best, have an intermediate truth value, between perfect truth and perfect falsity. For discussion see Williamson (1994, chaps. 4, 5 and 7). For an approach on which the soritical principle is maintained (and the paradox resolved by surrendering the transitivity of logical consequence) see Zardini (2008).
12. In discussing the JUSTIFIED POSSIBLE FUTURE UNLIKELYHOOD principle in n4, I suggested that, if a leaf has not fallen from the tree by the beginning of the *n*th hour there is a substantial objective probability that it will not have fallen from the tree by the end of the *n*th hour. If we accept this supposition, and reject JUSTIFIED AUTUMN LEAF, then we are obliged to reject JUSTIFIED POSSIBLE FUTURE UNLIKELYHOOD as well. The normic translation of this principle is as follows:

NORMIC POSSIBLE FUTURE UNLIKELYHOOD.

If it is normal, given my evidence, that *P* has or will have a substantial objective probability of being false, then it is normal, given my evidence, for *P* to be false.

This principle may have some superficial appeal, but is difficult to assess at a glance. Given suitable suppositions about the behaviour of objective probabilities, it will allow us to derive the soritical NORMIC ARRIVAL TIME and perhaps even NORMIC ALARM, which, I've argued, is subject to a clear counterexample. In any case, it is apparent that the appeal of NORMIC FAIR COINS owes nothing to this principle, and can be motivated directly by the present conception of normalcy or, arguably, by an analysis of 'fairness', along with certain formal principles. There is no tension in maintaining NORMIC FAIR COINS, while rejecting NORMIC POSSIBLE FUTURE UNLIKELYHOOD and thus, on a normic theory, there is no tension in maintaining JUSTIFIED FAIR COINS while rejecting JUSTIFIED POSSIBLE FUTURE UNLIKELYHOOD.

13. Some have claimed that many of the macrophysical events that we classify as normal or abnormal are in fact driven by microphysical processes that are analogous to a series of coin flips (Dutant, 2016, Section 2.2, see also Goodman, 2013, Section 3). Dutant (2016, Section 2.2) considers the diffusion of milk in tea and writes that the process 'involves many small collisions such that, if they all unfolded in some specific way, the milk could stay lumped together in the centre of the liquid instead of spreading more or less evenly. And every collision has some chance of going the requisite way' (Dutant 2016, 157). On close inspection, I think it is not so clear that this process, even so described, is relevantly analogous to a series of coin flips – not *fair* and *independent* coin flips at any rate. And, without something to play the role of the fairness and independence assumptions, the mere fact that a process has a non-zero objective probability of yielding a particular outcome does not imply that the outcome

should be regarded as maximally normal. A full treatment of this topic is beyond the scope of this paper, however. In any case, the normic theory does predict that judgments about what we have justification for believing will be hostage to judgments about what requires explanation: In so far as we suppose that a striking outcome, like the 'milk bubble' possibility, would require special explanation, we should take ourselves to have justification for believing that it won't happen. In so far as we suppose that a striking outcome would not require special explanation, and could occur purely as a matter of chance, the most we should take ourselves to have justification for believing is that it is very unlikely to happen. It is not implausible that these judgments should align in this way.

Acknowledgements

This paper was presented at the University of St Andrews in October 2016. Thanks to all of those who participated on this occasion. Thanks also to two anonymous referees for this journal.

Funding

Work on this paper was supported by the Arts and Humanities Research Council [grant number AH/M009610/1].

Notes on contributor

Martin Smith is a lecturer in philosophy at the University of Edinburgh working primarily on logic, epistemology and the philosophy of law.

References

- Bacon, A. 2014. "Giving Your Knowledge Half a Chance." *Philosophical Studies* 171 (2): 373–397.
- Dorr, C., J. Goodman, and J. Hawthorne. 2014. "Knowing Against the Odds." *Philosophical Studies* 170 (2): 277–287.
- Dutant, J. 2016. "How to be an Infallibilist." *Philosophical Issues* 26 (1): 148–171.
- Goodman, J. 2013. "Inexact Knowledge Without Improbable Knowing." *Inquiry* 56 (1): 30–53.
- Goodman, J., and B. Salow. forthcoming. "Taking a Chance on KK." *Philosophical Studies*: 1–14.
- Hawthorne, J., and M. Lasonen-Aarnio. 2009. "Knowledge and Objective Chance." In *Williamson on Knowledge*, edited by P. Greenough and D. Pritchard, 92–108. Oxford: Oxford University Press.
- Smith, M. 2010. "What Else Justification Could be." *Noûs* 44 (1): 10–31.
- Smith, M. 2016. *Between Probability and Certainty*. Oxford: Oxford University Press.
- Smith, M. forthcoming. "Why Throwing 92 Heads in a Row is not Surprising." *Philosophers' Imprint*.
- Williamson, T. 1994. *Vagueness*. London: Routledge.
- Williamson, T. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.
- Zardini, E. 2008. "A Model of Tolerance." *Studia Logica* 90 (3): 337–368.

Appendix I. A model of normalcy

This model is based upon the idea that claims about what is normal, and about how events normally unfold, can be treated as involving quantification over possible worlds that stand in relations of comparative normalcy. I introduce a simple formal language **L** consisting of countably many sentential constants ($A_1, A_2, A_3 \dots$), the truth functional operators ($\wedge, \vee, \sim, \supset, \equiv$), two special operators ($\blacklozenge, \blacksquare \rightarrow$) and punctuation. The sentences of **L** are built up using the standard recursive clauses for operators and punctuation. The semantics for **L** is provided by models of the form $\langle W, \leq, I \rangle$ where W is a set of possible worlds, \leq is a relation on W reflecting the comparative normalcy of worlds, given background evidence, and I is an interpretation function mapping pairs of sentences of **L** and members of W to the truth values $\{1, 0\}$. The relation \leq is assumed to be reflexive, transitive and connected (a total preorder on W). The function I is subject to the following constraints:

For any sentences $\phi, \psi \in \mathbf{L}$ and world $w \in W$:

$I_w(\phi \wedge \psi) = 1$	iff	$I_w(\phi) = 1$ and $I_w(\psi) = 1$
$I_w(\phi \vee \psi) = 1$	iff	$I_w(\phi) = 1$ or $I_w(\psi) = 1$
$I_w(\sim\phi) = 1$	iff	$I_w(\phi) = 0$
$I_w(\phi \supset \psi) = 1$	iff	$I_w(\phi) = 0$ or $I_w(\psi) = 1$
$I_w(\phi \equiv \psi) = 1$	iff	$I_w(\phi) = 1$ and $I_w(\psi) = 1$ or $I_w(\phi) = 0$ and $I_w(\psi) = 0$
$I_w(\blacklozenge\phi) = 1$	iff	for every world $x \in W$, if $I_x(\phi) = 0$ there is a world $y \in W$ such that $x \leq y$ and $I_y(\phi) = 1$
$I_w(\phi \blacksquare \rightarrow \psi) = 1$	iff	there is a world $x \in W$ such that $I_x(\phi) = 1$ and $I_x(\psi) = 1$ and there is no world $y \in W$ such that $x \leq y$ and $I_y(\phi) = 1$ and $I_y(\psi) = 0$

Put less formally, $\blacklozenge\phi$ is true iff for every world at which ϕ is false, there is an equally or more normal world at which ϕ is true, and $\phi \blacksquare \rightarrow \psi$ is true iff there is a world at which ϕ and ψ are true and which is more normal than any world at which ϕ is true and ψ is false. If we assume that there are maximally normal worlds (worlds in W that are maximal with respect to \leq) then $\blacklozenge\phi$ is true iff ϕ is true at some maximally normal worlds. If we assume that, for any sentence ϕ that is true at some possible worlds, there are maximally normal worlds at which it is true (worlds in $I(\phi)$ that are maximal with respect to \leq), $\phi \blacksquare \rightarrow \psi$ is true iff ψ is true at all maximally normal worlds at which ϕ is true. \blacklozenge might be read ‘It is normal that ...’ and $\blacksquare \rightarrow$ might be read ‘If ... it would normally be the case that ___’.

As can be checked, the following are valid inferences:

$\frac{\phi \blacksquare \rightarrow \psi}{\blacklozenge\phi \supset \blacklozenge\psi}$	$\frac{\blacklozenge(\phi \vee \psi)}{\blacklozenge\phi \vee \blacklozenge\psi}$
--	--

Let ϕ be the sentence ‘Coin c is flipped’, ψ be the sentence ‘Coin c is flipped and lands heads’ and χ be the sentence ‘Coin c is flipped and lands tails’. Let the background evidence include the fact that coin c is fair and no further relevant facts. Given this evidence, if coin c is flipped, it would normally land heads or tails – $\phi \blacksquare \rightarrow (\psi \vee \chi)$. This entails $\blacklozenge\phi \supset \blacklozenge(\psi \vee \chi)$ – if it is normal for coin c to be flipped then it is normal for coin c to land heads or tails. But $\blacklozenge(\psi \vee \chi)$ entails $\blacklozenge\psi \vee \blacklozenge\chi$ from which we derive $\blacklozenge\phi \supset (\blacklozenge\psi \vee \blacklozenge\chi)$. Given that the background evidence includes the fact that the coin is fair, we also have it that $\blacklozenge\psi \equiv \blacklozenge\chi$ – it is normal for the coin to be flipped and land heads just in case it is normal for the coin to be flipped and land tails. This is a special case of the conjecture flagged in the above discussion of NORMIC FAIR COINS: Given that a coin is fair, it would

be equally normal for it to be flipped and land heads and for it to be flipped and land tails. These premises together entail $\Diamond\phi \supset \Diamond\chi$ – if it is normal for coin c to be flipped, it is normal for coin c to land tails. This gives us the NORMIC FAIR COINS principle. **QED**

This model can be applied to Dorr, Goodman and Hawthorne's set-up as follows: For any n between 1 and 1000, let ϕ_n be the sentence 'Coin number n is flipped', ψ_n be the sentence 'Coin number n is flipped and lands heads' and χ_n be the sentence 'Coin number n is flipped and lands tails'. Let the background evidence include the coin flipping rules, the facts that coins 1 through to 1000 are all fair and no further relevant facts. Given the rules we have $\Diamond\phi_1$ and, for any n , $\Diamond\chi_n \supset \Diamond\phi_{n+1}$ – it's normal for the first coin to be flipped and, for any coin, if it is normal for the coin to be flipped and land tails, it's normal for the next coin in the sequence to be flipped. By the above result we have, for any n , $\Diamond\phi_n \supset \Diamond\chi_n$ – for any coin, if it is normal for the coin to be flipped then it is normal for the coin to be flipped and land tails. From this we can derive $\Diamond\phi_n \supset \Diamond\phi_{n+1}$ and, using the base case and this inductive step, we can prove $\Diamond\phi_{1000}$.

Goodman and Salow ([forthcoming](#), Section 5) apply a similar model to this set-up, but with the starting assumption that the outcomes vary according to their normalcy, with the abnormality of an outcome being a function of the number of consecutive tails that it involves. This could perhaps be underwritten by a picture on which the normalcy of a sequence of coin throws is associated with its 'randomness' or 'paternlessness' (see Dutant 2016, Section 2.3). From this starting point, we have it that, for some n , $\Diamond\phi_n$, $\Diamond\chi_n$ and $\Diamond\phi_{n+1}$ but $\sim\Diamond\chi_{n+1}$ – NORMIC FAIR COINS is violated. Since the model still validates $\Diamond\phi_{n+1} \supset (\Diamond\psi_{n+1} \vee \Diamond\chi_{n+1})$, we have $\Diamond\psi_{n+1}$. Although coin $n + 1$ is fair, it is normal for it to be flipped and land heads, but it is not normal for it to be flipped and land tails. As such, the model also violates the claim that, whenever a coin is fair, it would be equally normal for it to be flipped and land heads and for it to be flipped and land tails. One might treat this as a reductio of Goodman and Salow's starting assumption or, alternately, as an argument to the effect that the outcomes of a fair coin flip *can* differ with respect to their normalcy. In the latter case, NORMIC FAIR COINS would need to rest, after all, on certain substantial assumptions about normalcy. I won't pursue this further here.