


Taking a chance on KK

Jeremy Goodman¹ · Bernhard Salow² 

© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract Dorr et al. (Philos Stud 170:277–287, 2014) present a case that poses a challenge for a number of plausible principles about knowledge and objective chance. Implicit in their discussion is an interesting new argument against KK, the principle that anyone who knows p is in a position to know that they know p . We bring out this argument, and investigate possible responses for defenders of KK, establishing new connections between KK and various knowledge-chance principles.

Consider the following case from Dorr, Goodman and Hawthorne (2014), henceforth DGH:

Flipping Coins

1000 coins are laid out one after another: $C_1, C_2, \dots, C_{1000}$. A coin flipper will flip the coins in sequence until either one lands heads or they have all been flipped. Then he will flip no more. You know that this is the setup, and you know everything you are in a position to know about which coins will be flipped and how they will land. (DGH p. 278)

The case is interesting because, on pain of skepticism, it poses a counterexample to the following intuitive principle:

FAIR COINS: If you know that a coin is fair, and for all you know it will be flipped, then for all you know it will land tails.

✉ Bernhard Salow
bs416@cam.ac.uk

¹ USC School of Philosophy, Mudd Hall of Philosophy (MHP), Room 113, 3709 Trousdale Parkway, Los Angeles, CA 90089-0451, USA

² Trinity College Cambridge, Trinity Street, Cambridge CB2 1TQ, United Kingdom

For suppose that, as a matter of fact, the first coin will land heads. On pain of skepticism, you can know now (before the experiment) that the last coin won't be flipped.¹ So there must be a first coin, C_{n+1} , that you know won't be flipped. Since C_{n+1} is the *first* coin you know won't be flipped, it follows that, for all you know, C_n will be flipped. But you know that C_n won't land tails, since otherwise C_{n+1} would be flipped, and you know that that won't happen. So C_n is a counterexample to FAIR COINS.²

FAIR COINS follows from an attractive principle connecting knowledge and chance:

POSSIBLE FUTURE UNLIKELIHOOD: If for all you know, there is or will be a substantial objective chance that P, then for all you know, P.³

If FAIR COINS is false, so is POSSIBLE FUTURE UNLIKELIHOOD. But if they are both false, what explains their appeal?

DGH (p. 284–6) offer the following hypothesis. They maintain that, while POSSIBLE FUTURE UNLIKELIHOOD is false, a closely related principle may well be true:

ACTUAL FUTURE UNLIKELIHOOD: If there is or will be a substantial objective chance that P, then for all you know, P.

ACTUAL FUTURE UNLIKELIHOOD doesn't entail FAIR COINS. But it does entail

WEAK FAIR COINS: If a coin is fair and will be flipped, then for all you know it will land tails.

WEAK FAIR COINS is superficially quite close to FAIR COINS. In fact, it plausibly collapses into FAIR COINS given

KK: If you know that P, then you're in a position to know that you know that P.

For suppose you know that C_i won't land tails. By KK, you can know that you know this. Plausibly, you can also know WEAK FAIR COINS. But then you can deduce that C_i won't be flipped from the known facts (i) that you know that C_i won't land tails and (ii) that WEAK FAIR COINS is true. This would seem to allow you to know that C_i won't be flipped. So C_i is no counterexample to FAIR COINS. Since C_i was chosen arbitrarily, this suffices to establish FAIR COINS. Parallel reasoning can be used to argue that, given KK, ACTUAL FUTURE UNLIKELIHOOD collapses into POSSIBLE FUTURE UNLIKELIHOOD.

¹ DGH argue that the threat of skepticism cannot be confined to artificial cases like **Flipping Coins**.

² The claim that there is a first coin that you know won't be flipped relies on the law of excluded middle, which some may wish to reject here on account of vagueness as regards which coins you know won't be flipped. However, as DGH (footnote 5) point out, although giving up excluded middle might allow one to avoid accepting the negation of FAIR COINS, it does not in any obvious way allow one to actually accept FAIR COINS (on pain of skepticism).

³ Following DGH, we use 'substantial' to mean non-negligible; in this sense, chances less than .5 can still be substantial. DGH point out that this principle, and those mentioned later, may need to be qualified to accommodate knowledge under 'cheesy' modes of presentation, and cases of clairvoyance or involving time travel (if such cases are possible). In what follows we will suppress such qualifications.

Reasoning with the KK principle is extremely natural, so if the principle is ultimately false, then the above facts constitute a compelling explanation of our naive judgements about FAIR COINS and POSSIBLE FUTURE UNLIKELIHOOD. But they can also be seen as an interesting new argument *against* KK. After all, WEAK FAIR COINS and ACTUAL FUTURE UNLIKELIHOOD are highly plausible. If we have to choose between KK, these principles, and skepticism about the future, KK may start to look like the weakest link.

The aim of this paper is to evaluate this new argument against KK. In Sect. 1, we explain why it is importantly different from familiar objections to KK that rely on margin for error principles. In Sect. 2, we consider a KK-friendly treatment of the case; this treatment will provide a different account of why FAIR COINS and POSSIBLE FUTURE UNLIKELIHOOD are attractive despite being false. In Sect. 3, we show that the KK-friendly treatment cannot be generalized to a slight variant of the case; this observation allows us to argue that KK is incompatible with even a very weak version of the thought that substantial chance of falsity precludes knowledge. In sections 4 and 5 we argue that KK-enthusiasts can nevertheless accept a systematic non-skeptical treatment of cases like **Flipping Coins**. Section 6 concludes.

1 Chance versus margins for error

The above argument from ACTUAL FUTURE UNLIKELIHOOD and anti-skepticism about the future is importantly different from more familiar arguments against KK. In particular, it is distinct from Williamson's (2000, chapter 5) influential argument, which relies on the following margin for error principle: you cannot know P unless your belief that P could not easily have been mistaken. In cases where you couldn't easily have failed to believe as you do, this principle collapses to the claim that you cannot know that P unless P is true in all relevantly nearby possibilities. This principle is clearly in tension with KK, since knowing that one knows requires not just that P be true in nearby possibilities, but also that it be true in possibilities nearby any that are themselves nearby, which is a strictly stronger requirement, given that the relevant nearness relation is not transitive. By contrast, ACTUAL FUTURE UNLIKELIHOOD does not rely on such a margin for error principle. Instead, it is naturally seen as a generalization of the fact that knowledge implies truth—it is a way of formalizing the idea that we cannot know P when P has not yet been settled, in that P still has (or will have) a substantial chance of being false. That this gripping idea turns out to be inconsistent with KK is thus a new and powerful reason to be worried about KK.

Another disanalogy between the two arguments is that Williamson's margin for error principle concerns the modal status of one's belief, whereas ACTUAL FUTURE UNLIKELIHOOD concerns instead the status of the proposition believed. Indeed, Williamson does not himself accept ACTUAL FUTURE UNLIKELIHOOD (or even the weaker principles ACTUAL UNLIKELIHOOD and KNOWN UNLIKELIHOOD discussed below) since he thinks we can know many propositions about the future that have probabilistically independent and non-trivial chances, and that by deducing their conjunction from them we can thereby come to know that conjunction even if it has

(and we know it has) a substantial objective chance of being false.⁴ Insofar as this radical conclusion is an outgrowth of the margin for error framework, DGH are already rejecting that framework in offering ACTUAL FUTURE UNLIKELYHOOD as part of the explanation of our pretheoretical judgements concerning FAIR COINS.

2 Saving KK through Defeat?

What should an advocate of KK say in response to the argument from ACTUAL FUTURE UNLIKELYHOOD? This section explores the possibility of rejecting that principle in favor of

ACTUAL UNLIKELYHOOD: If there is a substantial objective chance that P, then for all you know, P.

This principle is slightly weaker, since a proposition can start out not having a substantial objective chance of being false, and yet come to have one in the future. As an extreme case, consider a world in which all 1000 coins are tossed and C_{1000} is the first to land heads. Then, at the start of the experiment, the objective chance that C_{1000} will (be tossed and) land tails is a miniscule (and so non-substantial) $.5^{1000}$. But by the time C_{999} has landed, the objective chance that C_{1000} will land tails will have risen to a very substantial .5. ACTUAL FUTURE UNLIKELYHOOD thus entails that you could not have known that C_{1000} won't land tails even at the start of the experiment; ACTUAL UNLIKELYHOOD has no such consequence.

At first sight, this might look like a reason *not* to give up ACTUAL FUTURE UNLIKELYHOOD for the weaker ACTUAL UNLIKELYHOOD. DGH (p. 286) comment that “it simply strikes us as implausible that in a world where C_{1000} is tossed and lands heads, we can know in advance that it won't land tails.”⁵ And if ACTUAL UNLIKELYHOOD were the only principle connecting knowledge and chance, it would be mysterious why someone living in a world like the one imagined above couldn't know, at the start of the experiment, that C_{1000} won't land tails.

However, things are not quite as bad as DGH make it seem. What we most immediately recoil from is the thought that someone could, once we reach C_{1000} , know that this coin, which is about to be tossed, won't land tails. And ACTUAL UNLIKELYHOOD has no trouble explaining why that can't happen: once we reach C_{1000} (if we do), there will be a substantial chance that it will land tails. What ACTUAL UNLIKELYHOOD doesn't predict is that, when the experiment is only just beginning, we already don't know that C_{1000} (which will in fact end up being tossed) won't land tails. But once we appreciate that we can't have a strong principle like FAIR COINS, it

⁴ See Williamson (2009). Although in what follows we ourselves will sometimes presume that one's knowledge can be extended by deductive inference, in no case will this involve more than one premise with non-trivial chance.

⁵ DGH bolster this judgment with a Williamsonian margin for error argument. But doing so amounts to giving up on the argument from ACTUAL FUTURE UNLIKELYHOOD as an independent and novel objection to KK. So we will set aside that argument here.

isn't clear that this prediction is something that any adequate response to the puzzle must vindicate.

Moreover, friends of KK have a principled general reason to be wary of ACTUAL FUTURE UNLIKELYHOOD. For almost any event E that presently has a very low nonzero chance of occurring, there is some chain of events, each of which has, given the ones before it, a reasonable chance of occurring, that would together make E quite likely. That is, almost all events that are presently unlikely are nonetheless likely to at some time in the future be likely to at some time in the future be likely to ... be likely. ACTUAL FUTURE UNLIKELYHOOD thus predicts that we cannot have iterated knowledge that the unlikely event won't occur, and KK collapses such ignorance into skepticism. By contrast, ACTUAL UNLIKELYHOOD does not seem to have such skeptical consequences given KK. After all, it looks plausible (at least at first sight) that the present chances always have chance 1 of having the values that they do.⁶ If that is right, then we cannot set up non-trivial iterations of claims about the present chances that would force us to choose between KK and skepticism.

The KK-friendly response we have been sketching comes with its own account of why ACTUAL FUTURE UNLIKELYHOOD (and hence, given KK, POSSIBLE FUTURE UNLIKELYHOOD) seems appealing: it follows from the true ACTUAL UNLIKELYHOOD if we don't take account of the fact that knowledge can be lost. In particular, suppose we accepted

NO DEFEAT: If you know P at the beginning of the experiment, then you know P throughout the experiment.

Given this principle, WEAK FAIR COINS follows from ACTUAL UNLIKELYHOOD. But reflection on the case suggests that NO DEFEAT is not beyond question. Perhaps you start off knowing that C_{1000} won't (be flipped and) land tails but this knowledge doesn't survive your seeing the first 999 coins land tails.

Embracing defeat (i.e., the possibility of knowledge being lost) to escape the slide towards POSSIBLE FUTURE UNLIKELYHOOD is a stable position. It comes with a new principle to approximate FAIR COINS:

VERY WEAK FAIR COINS: If a coin is fair and is about to be flipped, then for all you know it will land tails.

We can see that the resulting view is consistent by considering the following simple Kripke model. We represent worlds by positive integers, with each integer n representing the world in which the n th coin lands heads.⁷ We also represent times by positive integers, with t representing the time just before the t th coin is flipped (if it is). Letting m be the largest real number such that a $.5^m$ chance still qualifies as substantial, we can define the accessibility relation R_t for knowledge at time t as follows, where $R_t(n) =_{df} \{x : nR_t x\}$:

⁶ Note, however, that this principle is in tension with Humean theories of objective chance; see Lewis (1994).

⁷ For simplicity we suppose the experiment involves infinitely many coins and we ignore the possibility that they all land tails.

$$R_t(n) = \begin{cases} \{n\} & \text{if } n < t, \\ \{x : t \leq x \leq t + m\} & \text{if } t \leq n \leq t + m, \\ \{x : t \leq x \leq n\} & \text{if } n > t + m. \end{cases}$$

In words, $R_t(n)$ is the set of integers x such that, for all you know at time t in the world in which the n th coin will in fact land heads, the x th coin will land heads. The model thus describes your knowledge as follows. If the experiment has finished, you know the outcome. If the experiment is still ongoing, then an outcome is compatible with what you know as long as it involves either only the next m coins landing tails, or else some number of coins landing tails that is not greater than the actual number of subsequent tails, whichever is greater.

It is easy to verify that, as defined, R_t is transitive, so the model validates KK. Moreover, in the world where C_n lands heads, the model says that, at all times t such that $n - m \leq t \leq n$, C_n lands tails for all you know at t , thus vindicating ACTUAL UNLIKELIHOOD and VERY WEAK FAIR COINS (with ‘for all you know ...’ understood as ‘you do not know that not ...’).

3 The Problem with Defeat

Denying NO DEFEAT is an attractive way for the defender of KK to handle the version of **Flipping Coins** we have been imagining: the case in which you are watching the experiment unfold. However, the response breaks down in an only slightly different version of the case in which you don’t observe and aren’t told about the outcomes of the flips.

The problem we want to focus on is not that denying NO DEFEAT is less plausible for this alternative scenario. To be sure, if there is defeat in such cases, it cannot be of the familiar kind in which you lose your knowledge that P because you get new evidence against P, since you won’t learn about the initial coins coming up tails, and thus will never receive any evidence to suggest that C_{n+1} is likely to be tossed after all. But there might, for all we say here, be other kinds of defeat, which don’t require encountering misleading evidence. (The subject of Harman’s (1973, p. 143–4) dead dictator case might be a precedent: he too is supposed to lose his knowledge that the dictator has died despite never himself encountering the prevalent misleading evidence to the contrary.)

The problem we want to press is rather that any account of the version of **Flipping Coins** in which you never observe and are never told about the outcome of the coin flips should respect an additional constraint:

SETTLEDNESS: For all times t and t' after all the coins that will be flipped have been flipped, you know at t that C_i won’t be flipped if and only if you know at t' that C_i won’t be flipped.

This constraint is extremely plausible. For consider what it would take for it to fail. On the one hand, you might, some time after the experiment is over, suddenly gain new knowledge about which coins were flipped. But, since it was stipulated that you don’t observe any of the outcomes, and aren’t told about them either, such

knowledge would amount to a kind of clairvoyance. On the other hand, you might, some time after the experiment is over, lose some of your knowledge about its outcome. But why would that be? You do not receive new evidence about the experiment. We can stipulate that, since the experiment is over and no longer something others are interested in, no new evidence concerning it is coming into existence. There are no more changes to the chances of any of the relevant propositions. Loss of knowledge at this point would thus be completely inexplicable.

Surprisingly, *SETTLEDNESS* is incompatible with *KK*, anti-skepticism, and knowledge of *ACTUAL UNLIKELIHOOD*. Consider the case in which, as a matter of fact, C_1 lands heads, though you don't observe and aren't told about this or any other aspect of the outcome. By anti-skepticism, there is a first coin, call it C_{m+1} , such that you know, right after the first coin has been flipped (call this time t), that C_{m+1} won't be flipped. (We assume that $m > 1$.) So right after the first coin has been flipped, you don't know that C_m won't be flipped, though you do know that it won't land tails. Let t' be the time just before C_m would have been flipped if none of the other coins had landed heads first. By *ACTUAL UNLIKELIHOOD*, if you know at t' that C_m won't land tails, that can only be if C_m won't be flipped; since you know *ACTUAL UNLIKELIHOOD*, you will know this conditional at t' . By *SETTLEDNESS*, you will retain at t' your knowledge at t that C_m won't land tails. By *KK*, you will know at t' that you know at t' that C_m won't land tails. By putting this knowledge together with your knowledge of the aforementioned conditional, you are in a position to know at t' that C_m won't be flipped. Since you didn't know this at t (since C_{m+1} is the first coin that you then knew wouldn't be flipped), despite the experiment being over at t , we have a violation of *SETTLEDNESS*.

This argument illustrates a more general tension between *KK* and any kind of defeat for which (i) you don't know beforehand that it won't occur, regardless of whether it does, and (ii) you have no independent way of verifying whether it has occurred. For suppose that the defeat does not in fact occur. Then you continue to know. *KK* then allows you to know that defeat didn't occur, even though you have no independent way of verifying this. Maybe there are cases in which such knowledge acquisition occurs, but in the case of **Flipping Coins** it is clear that we should reject such seeming clairvoyance.

4 Doing Without Actual Unlikelihood

After some initial optimism about reconciling *KK* with *ACTUAL UNLIKELIHOOD* in Sect. 2, we saw in Sect. 3 that the two really are incompatible. *KK* enthusiasts must thus reject *VERY WEAK FAIR COINS*, meaning that they are, as yet, left with no account of why principles like *FAIR COINS* seem so compelling.

Can such an account be provided? Perhaps it can. For we could maintain that knowledge is lost in the version of **Flipping Coins** in which you are around to observe the outcomes, but not in the variant in which you are not informed about the outcomes. This proposal is compatible with the following yet weaker principle about the setup:

EXTREMELY WEAK FAIR COINS: If a coin is fair and you know that it is about to be flipped, then for all you know it will land tails.

That principle, like the previous ones, is an instance of a more general principle connecting knowledge and chance, namely

KNOWN UNLIKELYHOOD: If you know that there is a substantial chance that P , then for all you know P .

EXTREMELY WEAK FAIR COINS together with **NO DEFEAT** implies the instances of **WEAK FAIR COINS** that concern cases where you find out beforehand whether a coin will be tossed, as in the variant of **Flipping Coins** in which you are around to observe the outcomes. More generally, when we overlook the possibility that knowledge can be destroyed, we are likely to fall for the principle

MORE IS BETTER: If I have observed, or been told, everything about a coin that you've observed or been told about it, then I can know everything about how it will land that you can know.

which when combined with **EXTREMELY WEAK FAIR COINS** yields **WEAK FAIR COINS** in full generality. For if a coin will be tossed, someone could, presumably, observe or be told that it will be tossed. By **EXTREMELY WEAK FAIR COINS**, that person won't know that the coin won't land tails. So, by **MORE IS BETTER**, you can't know that the coin won't land tails either, in which case **WEAK FAIR COINS** is true.

On the resulting picture, the appeal of principles such as **WEAK FAIR COINS** (and hence **FAIR COINS**) ultimately rests on a hypothesized tendency to deny knowledge when we are aware of defeating evidence that the subject of knowledge is not aware of. For example, if I have conflicting reports from two different newspapers, and you have read only one of the newspapers, I may be tempted to deny that you know the fact that it reports. After all, any knowledge that you have would be of no use to me, since it would not survive once we started pooling our information. Nevertheless, this tendency to deny knowledge seems to us a mistake. If we both read a reliable newspaper, and later I unluckily come across a different newspaper with a misprint contradicting what the first newspaper reported, then I really do know less than you do despite having read more about the relevant subject matter.

The **KK**-friendly view we have been exploring thus seems to have the resources to explain the appeal of principles like **WEAK FAIR COINS**. But it comes at a cost that the earlier account in Sect. 2, if it had been general enough, would have managed to avoid. This cost arises from the fact that **KNOWN UNLIKELYHOOD** is quite different in structure from either **ACTUAL UNLIKELYHOOD** or **ACTUAL FUTURE UNLIKELYHOOD**: unlike the latter two principles, it does not express a direct connection between a proposition's chance and its knowability. A theory which vindicates only **KNOWN UNLIKELYHOOD** must thus reject the plausible thought, elaborated in Sect. 1, that chance of falsity is itself a barrier to knowledge.

This cost also highlights an explanatory challenge faced by the kind of view just sketched. On such a view, knowledge is compatible with substantial chance of falsity, but not with knowledge of substantial chance of falsity. But why would that

be? If you can know something that has a substantial chance of falsity, why can't you retain that knowledge when you find out about the chances?

We grant that this explanatory challenge has some pull, but we think that it can be met, at least in the case at hand. In the next section we will do so by subsuming the present account of **Flipping Coins** under a general and independently attractive account of knowledge.

5 Modelling KK and Defeat

The basic picture is this.⁸ We assume that people have a certain body of indefeasible evidence, and that they always know what this evidence is.⁹ In the case at hand, we will assume for simplicity that this includes all and only the facts about the setup and the outcomes that you are either told about or observe for yourself. Obviously, people know more than what is directly entailed by such Cartesian evidence. One way to capture this fact is by appeal to the thought that different possibilities vary in how normal or typical they are, and that we have a default entitlement to assume that things are relatively normal. We can thus reasonably believe that the state of the world is normal to the extent that this is compatible with our evidence. And, in so far as these beliefs are reliable, in that they aren't false in situations at least as normal as the ones we actually find ourselves in, they amount to knowledge.

However—and this is the crucial feature we add to the familiar picture—this default entitlement has only limited force. In particular, if a scenario is only slightly less normal than the most normal scenarios compatible with your evidence, then you are not entitled to ignore it simply on that basis. It is this feature which will ultimately allow for knowledge to be destroyed. Your default entitlement may initially justify you in putting a lower bound on how abnormal your situation is, and this lower bound may be correct so that this justified belief amounts to knowledge. Yet you may lose the right to insist on this lower bound if you then get evidence that rules out the most normal of your earlier evidential possibilities.

This picture suggests the following general formal account. Let a *normality structure* be a triple $\langle S, \leq, \ll \rangle$ where S is a non-empty set, \leq is a reflexive transitive relation on S , and \ll is a well-founded relation on S such that (i) if $s \ll s'$, then $s \leq s'$, and (ii) if $s' \leq s$, $s \ll t$, and $t \leq t'$, then $s' \ll t'$. Think of S as a set of *states*, read ' $s \leq t$ ' as *s is at least as normal as t*, and read ' $s \ll t$ ' as *s is far more normal than t*. We can model a body of evidence E as a non-empty subset of S , the set of the states not ruled out by that evidence.

⁸ Our picture is broadly inspired by 'normal conditions' approaches to knowledge. Greco's (2014) and Salnaker's (2015) defenses of KK against objections arising from considerations of reliability or margins for error are examples of such approaches. Goodman (2013, Sect. 3) argues that, whatever one thinks about margin for error principles, there is a distinct normality-theoretic condition on knowledge. Smith (2010, 2016) uses a related normal conditions idea to articulate a notion of justified belief. We plan to explore these ideas at greater length in future work.

⁹ We don't have any strong attachment to the word 'evidence'; those like Williamson (2000) who think that all knowledge is evidence should interpret us as proposing that there is some interesting subclass of one's evidence that plays the theoretical role we are about to describe.

We want to formalize the intuitive idea that we have a default entitlement to ignore possibilities that are much less normal than things might be given our evidence; that is, if we have evidence E , we can legitimately ignore worlds that are much less normal than some worlds compatible with E . Turning this around, the worlds compatible with what we are entitled to believe are just those members of E that are not much less normal than any member of E . This yields the following formal definition of $R_E^B(s)$, the set representing which states are compatible with what one can justifiably believe when one is in state s and has evidence E :

$$R_E^B(s) = \{s' \in E : \neg \exists s'' \in E, s'' \ll s'\}.$$

Note that $R_E^B(s)$ is independent of s ; this is because what one has justification to believe doesn't depend on what state the world is actually in, but only on what evidence one has.

Next, we want to move from this account of justified belief to an account of knowledge. The intuitive thought is that a justified belief amounts to knowledge just in case it is reliably formed, where a belief is reliable just in case it is true in all states compatible with one's evidence that are at least as normal as how things actually are. This thought naturally suggests the following definition of $R_E^K(s)$, the set of states compatible with what you are in a position to know when you are in state s and have evidence E :

$$R_E^K(s) = R_E^B(s) \cup \{s' \in E : s' \leq s\}.$$

In words: the states consistent with what you know when you're in state s and have evidence E are all the states which E doesn't justify you in disbelieving, together with all states that are consistent with your evidence and at least as normal as s .

Clearly, R_E^K is transitive; so, as desired, the model validates KK.¹⁰ And, as promised, the model also predicts that we sometimes lose knowledge as we gain evidence. Suppose, for example, that $S = \{1, 2, 3\}$, where $1 < 2 < 3$ and $1 \ll 3$ but $1 \not\ll 2$ and $2 \not\ll 3$. Suppose that 2 is the actual state. Given trivial evidence, one can then know that one isn't in 3: that belief is justified ($3 \notin R_{\{1,2,3\}}^B(2)$) because $1 \ll 3$) and reliable (being true at both 1 and 2, which are the only states ≤ 2). But if this evidence is augmented to rule out 1, one loses this knowledge, for the evidence now says that the world is at least somewhat abnormal, and so the default entitlement is no longer enough to justify ignoring 3. Or, as the formalism has it, $3 \in R_{\{2,3\}}^K(2)$ because $3 \in R_{\{2,3\}}^B(2)$, which in turn holds because no state in $\{2, 3\}$ is $\ll 3$.

¹⁰ Formally, we generate a Kripke model from a normality structure by identifying points of evaluation with state-evidence pairs $\langle s, E \rangle$ such that $s \in E$, where $\langle s', E' \rangle$ is compatible with what you are in a position to know at $\langle s, E \rangle$ if and only if $E = E'$ and $s' \in R_E^K(s)$, and likewise for belief and R^B . It can be shown the logic of knowledge and belief corresponding to the above clauses is the same as the one advocated by Stalnaker (2006), making the logic of knowledge S4.2. (A proof of this fact is beyond the scope of this paper.) Note that we can also use normality structures to model a margin-for-error requirement on knowledge, by instead defining $R_E^K(s)$ to be $R_E^B(s) \cup \{s' \in E : s' \leq s \text{ or } (s \leq s' \text{ and } s \not\ll s')\}$, which invalidates KK. So while the normality picture can be used to defend KK, it does not inexorably lead to it.

Here is how the general model applies to **Flipping Coins**. As before, we represent states by positive integers, with n representing the state in which the n th coin lands heads. We identify \leq with the less-than-or-equal-to relation on the positive integers. State n is far more normal than state n' if, in n , there is never a substantial chance of the sequence continuing at least as long as it does in n' ; that is, letting m again be the largest real number such that $.5^m$ is still a substantial chance, $\ll = \{ \langle n, n' \rangle : n + m < n' \}$.

Consider first the case where your evidence is trivial (i.e. $E = \{1, 2, 3, \dots\}$), as it is at the start of the experiment. We then get:

$$R_{\{1,2,3,\dots\}}^K(n) = \begin{cases} \{x : x \leq m + 1\} & \text{if } n \leq m + 1, \\ \{x : x \leq n\} & \text{if } n > m + 1. \end{cases}$$

Since $.5$ is a substantial chance, it follows that, unless the first coin lands heads, ACTUAL FUTURE UNLIKELYHOOD will be violated. In the version of the case where you get no new evidence, there will be no knowledge lost, so ACTUAL UNLIKELYHOOD will be violated too. If $n \leq m + 1$, this will happen immediately after the first coin lands tails, at which point the proposition that C_x will be flipped will come to have a substantial chance but will still be known to be false, where x is the least integer greater than $m + 1$. (In the substantially improbable event that $n > m + 1$, this won't happen until C_x lands tails—where x is the least integer greater than $n - m$ —at which point the proposition that C_{n+1} will be flipped will come to have a substantial chance but will still be known to be false.)

If your evidence is non-trivial, the model is more interesting. In particular, we can consider what you know at the various times t in the original version of **Flipping Coins**, in which you watch the experiment. In that case, your evidence $E(t)$ will be $\{n\}$ if $n < t$ and $\{x : x \geq t\}$ otherwise, which delivers exactly the results described in Sect. 2:

$$R_{E(t)}^K(n) = \begin{cases} \{n\} & \text{if } n < t, \\ \{x : t \leq x \leq t + m\} & \text{if } t \leq n \leq t + m, \\ \{x : t \leq x \leq n\} & \text{if } n > t + m. \end{cases}$$

The model thus predicts KNOWN UNLIKELYHOOD in **Flipping Coins**; and in doing so, it shows that there is, after all, a clear and coherent picture on which knowledge about the chances can, in this case, act as a defeater even though low chance is not itself a barrier to knowledge.

Of course, the fact that a natural normality model of this case vindicates KNOWN UNLIKELYHOOD does not show that the normality framework supports KNOWN UNLIKELYHOOD in full generality. To establish such a general connection, we would need a general principle linking normality and objective chance. Perhaps the most natural such principle is:

CHANCE-NORMALITY LINK: If in state s there ever is a substantial objective chance that P, then there is a state s' which is not far less normal than s in which P is true.

This principle gives voice to the intuitive idea that, if something could easily happen (in the sense of having a substantial chance of happening), then it happening would not be bizarre in comparison to whatever actually happens. And it entails KNOWN UNLIKELYHOOD, when combined with the assumptions (i) that your evidence has chance 1 and (ii) that the normality order is total and well-founded, so that every set has a (possibly non-unique) most normal element.¹¹ Unfortunately, it's not clear why (ii) would be true in general, even if it is plausible in particularly simple examples such as **Flipping Coins**: states can be abnormal in many different respects, and it is not clear that they can always be traded off against one another (though Smith's (2016, chap. 6–7) use of concentric spheres to model normality orders indicates that he thinks that they can be). And without (ii), there is no route from CHANCE-NORMALITY LINK to KNOWN UNLIKELYHOOD.¹²

Moreover, the normality picture is, at least superficially, in tension with KNOWN UNLIKELYHOOD, because normality models validate the closure of knowledge under conjunction. Imagine a setup involving a million independent copies of the **Flipping Coins** setup. Now consider the normality model of one's knowledge of the outcomes of these 'experiments' in which what you know about the outcome of one experiment depends only on the outcome of that experiment, in the way proposed above. This model seems compatible with the general normality picture. However, it is not compatible with KNOWN UNLIKELYHOOD (or even the weaker CHANCE-NORMALITY LINK). For let $\{f_1, \dots, f_n\}$ be the set of experiments (in all likelihood, the great majority of them) that will have normal outcomes. In the model, you will know before the experiments that f_1, \dots, f_n will all have normal outcomes. But given the setup, this proposition has a low objective chance, on account of the extremely large number of experiments being run. Since we may assume that you know this consequence of the setup, we have a counterexample to KNOWN UNLIKELYHOOD.

So much for the project of deriving KNOWN UNLIKELYHOOD from the normality picture and principles connecting normality and chance. But notice that, irrespective of whether KNOWN UNLIKELYHOOD can be justified in this way, more specific principles such as EXTREMELY WEAK FAIR COINS definitely can be. We need to assume only that (i) for any two possible world-histories that agree up to the point of some potential future event e , where e is the event of some particular fair coin being

¹¹ Suppose you justifiably believe that P has a substantial objective chance, i.e. P has a substantial objective chance throughout $R_E^B(s)$. You know that E is your evidence; so by (i) E has chance 1 in every state in $R_E^B(s)$. So E&P has the same chance as P, and hence a substantial chance, throughout $R_E^B(s)$. By (ii), E has at least one element at least as normal as all the others, call it t . Clearly, t is in $R_E^B(s)$, so E&P has a substantial chance at t . By CHANCE-NORMALITY LINK, there is a state t' such that $t \ll t'$ in which E&P is true. Since t was at least as normal as anything else in E, and $t \ll t'$, we have that $t'' \ll t'$ for all t'' in E. So $t' \in R_E^B(s)$, and so P is compatible with what you have justification to believe.

¹² To see why the assumption is essential, consider the model $S = \{1, 2, 3, 4\}$, with $1 \leq 3$, $1 \ll 3$, $2 \leq 4$, $2 \ll 4$, and no other normality relations between states. Then it is consistent with CHANCE-NORMALITY LINK that $\{3, 4\}$ has a substantial chance at both 1 and 2. But, in that case, you can, in 1 and 2, know both that $\{3, 4\}$ has a substantial chance, and that it's not true. (We haven't been able to think of an example that intuitively exhibits this structure; perhaps this is because there is a connection between chance and normality, which we've been unable to identify, which does guarantee that KNOWN UNLIKELYHOOD holds in all normality structures.)

flipped, your evidence is either compatible with both of these world-histories or with neither of them, and (ii) no possible world-history in which e obtains, where e is the event of some fair coin being flipped, is far more normal than every such history in which e results in a tails outcome (and vice versa). The first assumption articulates the thought that your evidence can speak decisively only to issues that have already been settled (which also motivated our earlier assumption that your evidence has chance 1); the second assumption is an alternative way of articulating the connection between chance and normality motivating the CHANCE-NORMALITY LINK. Together, these two principles entail EXTREMELY WEAK FAIR COINS when we enrich our normality models with extra structure for representing the settledness of such events, as can be straightforwardly done using the formalism of branching time. Generalizations of these principles will naturally entail that whenever you know that a particular chance event is about to occur and has a substantial chance of having a particular outcome, then for all you know it will have that outcome. And, unlike KNOWN UNLIKELYHOOD, *this* kind of knowledge-chance principle does not require us to choose between anti-skepticism and the closure of knowledge under conjunction (since a plurality of many potential future chancy events is not itself a potential future chancy event). So even advocates of KK and the normality picture who reject KNOWN UNLIKELYHOOD are left with a coherent and principled picture of the interaction between knowledge and chance.

6 Conclusion

Given anti-skepticism about the future, KK surprisingly turns out to be incompatible with ACTUAL UNLIKELYHOOD. Denying that principle means that we abandon the idea that chanciness is any sort of barrier to knowledge. And we do so in an unprecedented way. For unlike worries about ACTUAL UNLIKELYHOOD that arise from the principle that what one is in a position to know is closed under conjunction, maintaining KK requires rejecting the further plausible principle VERY WEAK FAIR COINS. And while KK is consistent with KNOWN UNLIKELYHOOD, their combination is in certain respects unattractive and the explanatory power of the latter principles is unclear.

On the other hand, natural ideas about knowledge and normality which entail KK lead to an elegant model of **Flipping Coins** in which KNOWN UNLIKELYHOOD holds even though ACTUAL UNLIKELYHOOD does not. Moreover, the appeal of VERY WEAK FAIR COINS can arguably be explained by the fact that it follows from EXTREMELY WEAK FAIR COINS together with principles that are intuitively tempting, but are nonetheless refuted by reflections on how knowledge can be destroyed. Perhaps, then, KK can live to see another day.

Acknowledgements We'd like to thank Andrew Bacon, Fabrizio Cariani, Catrin Campbell-Moore, Cian Dor, Kevin Dorst, Peter Fritz, John Hawthorne, Harvey Lederman, Jeff Russell, Bob Stalnaker, and an anonymous referee for *Philosophical Studies*, as well as audiences at NYU, the 2015 Puzzles of Knowledge workshop at the University of Lisbon, and the 2016 Formal Epistemology Workshop.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Dorr, C., Goodman, J., & Hawthorne, J. (2014). Knowing against the odds. *Philosophical Studies*, *170*, 277–287.
- Goodman, J. (2013). Inexact knowledge without improbable knowing. *Inquiry*, *56*, 30–53.
- Greco, D. (2014). Could KK be OK? *Journal of Philosophy*, *111*, 169–197.
- Harman, G. (1973). *Thought*. Princeton: Princeton University Press.
- Lewis, D. (1994). Humean supervenience debugged. *Mind*, *103*, 473–490.
- Smith, M. (2010). What else justification could be. *Nous*, *44*, 10–31.
- Smith, M. (2016). *Between probability and certainty: What justifies belief*. Oxford: Oxford University Press.
- Stalnaker, R. (2006). On logics of knowledge and belief. *Philosophical Studies*, *128*, 169–199.
- Stalnaker, R. (2015). Luminosity and the KK thesis. In S. Goldberg (Ed.), *Externalism, self-knowledge, and skepticism*. Cambridge: Cambridge University Press.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Williamson, T. (2009). Probability and danger. *The Amherst Lecture in Philosophy*, *4*, 1–35.