

A Paradox of Evidential Equivalence

David Builes

Forthcoming in *Mind*

Different pieces of evidence are about different things. Some of our evidence is about coins, and some of our evidence is about dinosaurs. It is natural to think that there is some interesting connection between facts about what a piece of evidence is about, *evidential aboutness*, and facts about how that piece of evidence bears on various hypotheses, *evidential relevance*. Here are a couple of examples. Suppose that I am about to roll a pair of dice. Any evidence that is entirely about how the first die will land seems to be evidentially irrelevant to how the second die will land. After all, the two dice are independent of one another. Now consider a more philosophical example, the Ravens paradox. Intuitively, the fact *that this shoe is white* seems to be evidentially irrelevant to the claim *that all ravens are black*. After all, the fact *that this shoe is white* is not at all about ravens. Of course, there may well be good reason to resist this intuitive verdict. At this point, these examples merely serve to illustrate a certain phenomenon.¹

Whatever the connection is between evidential aboutness and evidential relevance, it seems obvious that it should respect the following platitude about evidence:

Evidential Equivalence: If one is (rationally) certain that E_1 is true iff E_2 is true, then for any H , it is rationally required that $\text{Cr}(H \mid E_1) = \text{Cr}(H \mid E_2)$.²

In other words, even E_1 and E_2 differ over what they are about, they should surely have the same evidential relevance towards H if one is rationally certain that they are either both true or both false.³

The purpose of this paper is to present a paradox that seems to cast some doubt on *Evidential Equivalence*, by exploiting certain fine-grained features of evidential aboutness. While I ultimately wish to retain *Evidential Equivalence*, I believe the paradox shows that our intuitive conceptions of *inadmissible* evidence and *independent* evidence are sensitive to facts about evidential aboutness in an interesting way.

¹ For a book-length treatment of the phenomenon of aboutness and its connections to several other notions, see Yablo (2014). For a recent assessment of different theories of aboutness, see Hawke (2018).

² Throughout, I will write $\text{Cr}(H)$ for an agent's unconditional credence that H is true and $\text{Cr}(H \mid E)$ for an agent's conditional credence that H is true given E .

³ Krämer (2017) also discusses the relation between evidential aboutness and evidential relevance, but he does not question *Evidential Equivalence*.

In §§1 and 2, I analyze two cases and argue that we ought to have a particular conditional credence in those cases, and in §3 I show why these conditional credences violate *Evidential Equivalence*. In §4, I respond to some natural worries about the case, and in §§5 and 6 I give some independent motivations for and against *Evidential Equivalence*.

1. Finite Coins

Consider the following case:

Finite Coins: Suppose you are in a room with a countable infinity of people, and each of you flips a coin without looking at the result. You know that all of the coin flips are fair and independent. I then inform you that something remarkable happened: *almost every coin landed tails*. More precisely, only finitely many coins landed heads. Now what should your credence be that your coin landed heads?

It seems like your credence should drop from $1/2$. Given that almost every coin landed tails, what's the chance that you are one of the vanishingly few heads-flippers? After all, the evidence that *almost every* coin landed tails should at least count as *some* evidence that you are one of the tail-flippers. Analogously, if you have a ticket for a lottery where almost every ticket is a losing ticket (analogously, almost every coin is a tails-coin), then if the lottery is fair (if every coin is equally likely to be one of the heads-coins), then you should think you probably have a losing ticket (you probably have a tails-coin). Here are some more formal arguments for the stronger conclusion that you should in fact lower your credence to 0.⁴

First, there is an accuracy argument. For every agent A in the room, let H_A be the proposition that agent A flipped heads, and suppose you wanted to minimize your inaccuracy with respect to each H_A . On standard measures of inaccuracy, such as the Brier score⁵, if you stick with $1/2$ for each H_A upon being informed that only finitely many agents flipped heads, you will accrue an equal finite amount of inaccuracy for each H_A .⁶ Since there are infinitely many agents in total, you will expect to accrue an infinite amount of inaccuracy.⁷ However, suppose you drop your credence to

⁴ One might instead want to assign some infinitesimal probability rather than 0. This is fine – all of the arguments in this paper will go through if one replaces ‘0’ with some infinitesimal quantity. Strictly speaking, I will only need the claim that one’s credence should be lower than $1/2$ in §3. However, for general arguments against the use of infinitesimals, see Easwaran (2014).

⁵ For much more on measures of inaccuracy and their justifications, see Leitgeb and Pettigrew (2010).

⁶ I am assuming here (and in some of the other arguments below) that one should assign equal credence to each H_A by symmetry considerations. The evidence that only finitely many coins landed tails is entirely neutral on which coins were tails and which coins were heads.

⁷ For example, on the Brier score, your inaccuracy for each H_A will be $(1-1/2)^2 = (0-1/2)^2 = 1/4$. It is also worth noting that assigning *any* non-zero credence c to each of the H_A will result in one having an infinite

0 for each H_A . Then, you will accrue no inaccuracy for all the tail-flippers, and you will accrue some equal finite amount of inaccuracy for each head-flipper. Since there are only finitely many head-flippers, you know you will only have a finite amount of inaccuracy. Having a finite amount of inaccuracy is better than having an infinite amount, so you should drop to 0 in each H_A rather than sticking to $1/2$ in each H_A .

Second, there is a Dutch book argument. Suppose you stick to $1/2$ for each H_A . Then, there will be a series of Dutch books, each of which is strictly favorable to you, for which you are guaranteed to lose an infinite amount of money.⁸ For each agent A , you would agree to receive \$2 if A flips heads and pay \$1 if A flips tails. No matter what happens, you will lose an infinite amount of money. In contrast, if you drop to 0 for each H_A , there will be no Dutch book against you with respect to the H_A . First, note that for a bet on H_A to be strictly favorable to you it must be of the following form: pay \$ X if A flips heads and receive \$ Y if A flips tails, where X is any real number and Y is any positive real number. No collection of bets of this form will guarantee a loss, since you might win all the bets if all of the agents flip tails.^{9,10}

Third, consider what you should do if you were instead informed that exactly n coins landed heads for some fixed constant n . If there were n^*m people in the room in total, your credence should go to $1/m$. So, in the limit as m goes to infinity and the number of people in the room increase, your credence should limit to 0. So, I claim that if there were countably many people in the room who flipped fair coins, and you were informed that exactly n coins landed heads, your credence that your coin landed heads ought to drop to 0. For suppose you did not align your credence to the limiting value of 0, and instead you set your credence to $c > 0$ in the infinite case. Then, there will be some number M such that $1/M < c$. So this means that, if there were n^*M total people in the room and n people flipped heads, your credence that your coin landed heads would be *higher* than if there were infinitely many total people in the room! Surely that can't be right. So, if there were countably many people in the room who flipped fair coins and n people flipped heads, your credence that your coin landed heads ought to drop to 0. So, in the original case, upon being informed that finitely many coins landed heads, you know that for *some* n , there are exactly n people who flipped heads. You also know that *no matter what the value of n is*, conditional on its

amount of inaccuracy. For each tail-flipper A , your inaccuracy in H_A will be $(0-c)^2=c^2$. Since there are infinitely many tail-flippers, one will accrue an infinite amount of inaccuracy.

⁸ It has been argued persuasively by Easwaran (2013) that Dutch books in infinitary cases should consist of individual bets that are strictly favorable.

⁹ As with footnote 7, it should be noted that assigning *any* fixed non-zero credence to each of the H_A is vulnerable to this Dutch Book Argument.

¹⁰ It is worth noting that the Dutch book considered here has several good-making features that might ward off some skepticism about infinitary Dutch books. First, it is a synchronic Dutch book. Second, every bet is uniformly bounded above and below. Third, there is no possibility that an infinite amount of money is both gained and lost because of this Dutch book.

true value, you ought to drop to 0. Therefore, it seems that you ought to drop to 0 merely upon being informed that finitely many coins landed heads.¹¹

This argument goes hand in hand with a reflection argument and an argument from deferring to epistemic experts. For the reflection argument, suppose that, after you are told that only finitely many coins landed heads, an announcer says that he will say the exact number of head-flippers in 30 seconds. You know that, no matter what he says, your credence should drop to 0. So why bother waiting? For the deference argument, suppose you know that the announcer told Bob the exact number of head-flippers in the room. Suppose you also know Bob is a perfectly rational agent who has all the evidence you have. You therefore know that Bob's credence that your coin landed heads is 0, and you know that he is a rational agent who has strictly more relevant evidence than you do, so you should have the same credence as him. Intuitively, however, the addition of the announcer and Bob is irrelevant. Whatever your credence should be in the original case, it should be the same in the modified cases with the announcer and Bob.

2. Finite Coins*

Next, consider the following variant of the case:

Finite Coins*: Again, you are in a room with countably many people and each of you flips a coin without looking at the result. You know that all of the coin flips are fair and independent. This time, you will only be told information about the *other* people in the room, excluding you. Let S be the set of these other people. I inform you of the following remarkable piece of information: *only finitely many people in S flipped heads*. What should your credence be that your coin landed heads?

The answer to Finite Coins* seems obvious: you should clearly stay at 1/2! The piece of information you received has nothing to do with your coin, since (it is stipulated that) you are certain that each flip is independent of any other. Because of this, your rational credence in the

¹¹ The last step appeals to countable conglomerability. Formulated as a rational constraint, countable conglomerability is the thesis that for any countable partition of mutually exclusive and exhaustive events E_1, E_2, E_3, \dots , if $c_1 \leq \text{Cr}(P | E_i) \leq c_2$ for all i , then it is rationally required that $c_1 \leq \text{Cr}(P) \leq c_2$. Some people deny the general thesis of countable conglomerability. In response, I have three points. First, neither the intuitive argument in terms of the lottery-analogy, nor the accuracy argument, nor the Dutch book argument explicitly appeal to countable conglomerability as a premise. Second, even if one denies the general principle of countable conglomerability, one might still want to retain this particular instance of it. Third, the purpose of this section is to present the strongest case in favor of lowering one's credence from 1/2. While some deny the general thesis of countable conglomerability, many endorse it. In the next paragraph, I will intuitively motivate countable conglomerability in the standard way.

proposition that your coin landed heads should not change upon learning that only finitely many people in S flipped heads.

Consider the following modified case:

Past Coins: This time, you are alone in a room holding a fair coin that you have yet to flip. Before you flip your fair coin, someone informs you of the following fact: last year, there was once a time where countably many people in this same room all flipped coins, and, remarkably, only finitely many of them landed heads. After receiving this curious bit of information about the past events in the room, what should your credence be that the fair coin you haven't even flipped yet will land heads?

The answer seems clear – your credence should be $1/2$. This follows from the Principal Principle. You should conform your credence to the objective chance of $1/2$ that your unflipped fair coin will land Heads. Could one resist this conclusion by claiming that the evidence you received about the events that transpired last year counts as 'inadmissible' information? Given that the evidence you received was entirely about events in the past (and did not involve any exotic information about time travelers or crystal balls), this suggestion is implausible. Could one think that Past Coins should be treated differently than Finite Coins*? It's hard to see how the mere temporal distance between your flip and the other people's flips could be relevant. Surely it shouldn't be relevant if the other people flipped 1 minute earlier than you, rather than 1 year earlier than you. It's hard to believe that there's a crucial difference if the flips happened *simultaneously* with yours (as in Finite Coins*), rather than slightly before.

One potential asymmetry in the two cases is that in Past Coins you are given a particular qualitative property, namely *being temporally separated from all other coin-tossings*, which singles out your coin-toss from the others, but you are given no such qualitative property in Finite Coins*. One might then worry that Finite Coins* (and Finite Coins) involves 'essentially indexical' or 'self-locating' propositions. While I am skeptical that such a difference should matter, for my purposes we can simply side-step this issue by stipulating that you *do* have a qualitative way of picking yourself out among the coin-flippers in Finite Coins* (and Finite Coins). Perhaps, for example, you are certain that you are the only person in the room wearing a red shirt. The proposition in question, that you flipped heads, will then be equivalent to the purely qualitative proposition that the red-shirted person flipped heads.¹² Adding this extra stipulation to the description of Finite Coins and Finite Coins* does not, as far as I can see, affect the intuitive verdicts about these cases.

¹² Dorr (2010) presents an interesting case in which he argues that one's credence that a certain future coin toss will be Heads should be 1, conditional on the outcomes of certain past coin tosses. Dorr's puzzle, however, *does* seem to essentially involve certain qualitative temporal symmetries and self-locating propositions. For his case, Dorr suggests that the application of the Principal Principle might need to be restricted to apply only to propositions with certain 'modes of presentation'. However, Dorr says, 'Nothing I have said generates any obvious worry about the Principal Principle as applied to purely

3. The Paradox

Let E_1 be the original piece of evidence that finitely many people in total flipped heads. Let us generalize the second piece of evidence. Let $E_{2,A}$ say that, excluding agent A , finitely many people in the room flipped heads. Note that for every agent A , E_1 is known to be necessarily equivalent to $E_{2,A}$. In other words, necessarily, finitely many people in total flipped heads if and only if finitely many people, excluding A , flipped heads.

Let Cr stand for the credence function you ought to have. The following three propositions form an inconsistent triad:

- (1) For some A , $Cr(H_A | E_1) < 1/2$
- (2) For every A , $Cr(H_A | E_{2,A}) = 1/2$
- (3) For every A , $Cr(H_A | E_1) = Cr(H_A | E_{2,A})$

Note that I have argued for a much stronger version of proposition 1 in §1, namely that for all A , $Cr(H_A | E_1) = 0$. However, the much weaker version of the proposition suffices to generate the contradiction. Proposition 3 is simply an instance of *Evidential Equivalence*.

4. Infinitary Worries

In this section, I will respond to three natural infinitary worries.

First, one might worry that the coins can't be 'fair' given that almost all of them landed tails. This might be a valid worry if one had some sort of frequentist view of chances on which *what it is* for a coin to have a 50-50 chance of coming up heads *just is* for the actual (or hypothetical) frequencies of certain coin flips (relative to a certain reference class) to have a limiting frequency of 50-50. However, frequentist views of chance are widely considered to be implausible. For a total of 30 arguments against frequentist views of chance, see Hájek (1996, 2009).

Second, one might worry that all of the conditional credences in the inconsistent triad above should just be 'undefined' since according to the Ratio Formula, $Cr(A | B) = Cr(A \wedge B)/Cr(B)$. In our case, we are conditioning on a proposition that has probability 0, since $Cr(E_1) = Cr(E_{2,A}) = 0$, so applying the ratio formula to $Cr(H | E_1)$ and $Cr(H | E_{2,A})$ does not give us a well-defined result.¹³

qualitative propositions...' (p. 202). The puzzle I will be focusing on can be run on purely qualitative propositions.

¹³ An interesting question arises about whether one can construct a similar case against *Evidential Equivalence* that doesn't rely on the evidence having probability 0. In other words, can one find examples of H , E_1 , and E_2 such that one is certain that E_1 is true iff E_2 is true, and $Cr(E_1) = Cr(E_2) > 0$, yet $Cr(H | E_1)$

In response, it should be noted that many presentations of the Ratio Formula explicitly have a clause that the formula is only valid *when* $Cr(B) \neq 0$. Many authors have argued that the Ratio Formula is simply silent in cases where $Cr(B) = 0$. Hájek (2003) has argued at length that the concept of conditional credence is at least as fundamental as that of unconditional credence, and the Ratio Formula should not be treated as a definition of conditional credence, but rather as a thesis about conditional credence that is only valid in certain unproblematic cases (such as when $Cr(B)$ is not 0). Hájek argues for an account on which conditional probability is a primitive two-place function not defined in terms of unconditional probability at all. Unconditional probabilities are then defined as conditional probabilities conditional on the tautology. This sort of primitivist account is also described in Popper (1955) and Rényi (1970). Even bracketing this sort of account, it just seems clear that probabilities conditional on possible probability 0 events can be made perfect sense of. What's the probability I will flip infinitely many heads, conditional on me flipping infinitely many heads? Obviously 1! What's the probability I will only flip finitely many heads, conditional on me flipping infinitely many heads? Obviously 0! What's the probability that I will roll a '6' with a fair die, conditional on Bob flipping infinitely many heads? Obviously 1/6! Here's a less trivial example. Suppose I flip a coin infinitely many times. Conditional on my infinite string of coins landing either HHHHHHHH... or TTTTTTTTTT... or HTHTHTHTHTHT..., what should my credence be that my second toss was tails? It should be 2/3, since any infinite string is as likely as any other and in two out of the three possible strings my second toss is heads. These examples show that there are rational requirements on conditional credences where the condition is assigned probability 0. The three propositions in the inconsistent triad above are just claims about what some of these rational requirements are.

Third, one might worry that as finite agents we can never actually *learn* or *update* on a probability 0 event. For example, we might be *told* that only finitely many coins landed heads, but we would just suspect that that was a lie. In response, it should be noted that the inconsistent triad above is entirely in terms of conditional probabilities. Even if it is in principle impossible to update on a probability 0 event (which is very contentious), I do not need to assume that it is possible to run the paradox.¹⁴

Given these three points, what are the prospects for a view that denies all three propositions in the inconsistent triad by saying that all credences conditional on zero-probability propositions (henceforth, 'null-probabilities') are 'ill-defined'? I see three problems with such a view. First, it

is intuitively different than $Cr(H | E_2)$? I'm not sure if such intuitive examples exist. However, if such a case does exist, it will run afoul of the Ratio Formula. Given the Ratio Formula, $Cr(H | E_1) = Cr(H \wedge E_1) / Cr(E_1)$, and $Cr(H | E_2) = Cr(H \wedge E_2) / Cr(E_2)$. These two values will be equal since $Cr(H \wedge E_1) = Cr(H \wedge E_2)$ and $Cr(E_1) = Cr(E_2)$ (given that one is certain that E_1 is true iff E_2 is true (and hence $H \wedge E_1$ is true iff $H \wedge E_2$ is true)).

¹⁴ It should be noted, however, that two of the arguments in support of proposition 1, namely the reflection and deference arguments, do implicitly assume that it is at least in principle possible to update on such propositions.

seems like many null-probabilities are entirely unproblematic. Surely it is a rational requirement that $\text{Cr}(A | A) = 1$, given that A is any contingent proposition compatible with one's evidence. Second, it seems that scientific practice is committed to some null-probabilities. As Myrvold (2015) notes, it seems that we *have* to regard some null-probabilities as well-defined in order to do justice to statistical practice. Statistical practice uses likelihood functions that assign well-defined null-probabilities to data conditional on particular (probability zero) point values of some continuously varying parameter. Third, as Dorr (2010) argues, it seems that we need well-defined null-probabilities to give a satisfactory account of objective chance. Null-probabilities are needed to express how the chances at earlier times evolve into the chances at later times. The chance function at some later time t_2 is just the chance function at some earlier time t_1 conditional on the complete truth about history between t_1 and t_2 , whose chance at t_1 may well be 0. For these reasons, I regard the position that all null-probabilities are simply 'ill-defined' as far too radical.¹⁵

5. Hyperintensional Evidence

So far, I have only argued in favor of the first two propositions in the inconsistent triad above. In order to put the case against *Evidential Equivalence* in its strongest light, I would like to briefly give some independent, positive motivation for getting out of the paradox by rejecting proposition 3, and hence rejecting *Evidential Equivalence*. One way to reject *Evidential Equivalence* is to think that Bayesians should treat evidence *hyperintensionally* – by making important epistemic distinctions among necessarily equivalent pieces of evidence. Why think that evidence is hyperintensional? Because *aboutness* is hyperintensional, and it is natural to think that facts about evidential relevance are sensitive to what the evidence is about. In the inconsistent triad above, E_1 is evidentially relevant to the state of everyone's coin since it is partially about everyone's coin. On the other hand, $E_{2,A}$ is not evidentially relevant to the state of A 's coin since it is not even about A 's coin. I believe that the strong intuitions pulling us in opposite directions in the inconsistent triad are entirely due to the hyperintensionality of aboutness.

The cases of Finite Coins and Finite Coins* motivate the thought that certain central concepts in Bayesian epistemology, namely *independence* and *inadmissibility*, are intuitively

¹⁵ Both Easwaran (2008) and Myrvold (2015) argue for a view according to which null-probabilities have to be further relativized to some contextually salient partition to have a well-defined answer. The partition is then used to specify how to compute the null-probabilities as a certain limit of unconditional probabilities. On this view, perhaps proposition 1 is true relativized to one partition, proposition 2 is true relativized to a distinct partition, but proposition 3 is false according to either partition. This is an interesting and controversial view which can't be fully assessed here. The view does go against the orthodox view that conditional probability must be an absolute rather than a relativized notion. For example, Kadane et al (1986) say, 'This approach is unacceptable from the point of view of the statistician who, when given the information that [some event] has occurred, must determine the conditional distribution of X_2 ' (p. 70).

hyperintensional. Intuitively, we want to say that evidence that is entirely *about* the outcomes of distinct coin tosses should be regarded as *independent* of your own coin toss. This has the consequence that the evidence received in Finite Coins* should be regarded as probabilistically independent of your own coin toss, while the evidence received in Finite Coins need not be regarded as probabilistically independent.

Next, turn to the concept of *inadmissibility*. Recall the following formulation of the Principal Principle given by Lewis (1980):

Principal Principle: Let C be any reasonable initial credence function. Let t be any time. Let x be any real number in the unit interval. Let X be the proposition that the chance, at time t , of A 's holding equals x . Let E be any proposition compatible with X that is admissible at time t . Then $C(A \mid X \wedge E) = x$.

Different understandings of ‘admissible’ lead to different versions of the Principal Principle. Lewis gives us the following sufficient condition for admissibility: ‘if a proposition is entirely about matters of particular fact at times no later than t , then as a rule that proposition is admissible at t ’ (p. 272). Note the crucial word ‘about’. We can bring out the hyperintensionality of ‘about matters of particular fact at times no later than t ’ by using the case of Past Coins in §2. The information that you received in that case, namely that last year countably many people flipped coins in your room and only finitely many coins landed heads, is ‘entirely about matters of particular fact’ in the past. So, it should count as admissible with respect to the chance of the proposition that your coin will land heads. However, if we let S be the set of all coins flipped at that time together with your coin, the proposition that only finitely many of the coins in S landed heads does not seem to be ‘entirely about matters of particular fact’ in the past, even though it is equivalent to the information that you received in Past Coins.

In sum, one might try to independently motivate the rejection of *Evidential Equivalence* by reflecting on the fact that certain central concepts in Bayesian epistemology, namely *independence* and *inadmissibility*, are intuitively hyperintensional.

6. Evidential Equivalence

Given the above considerations, one might think that the right way to respond to the paradox is simply to deny proposition 3 by denying *Evidential Equivalence*. However, rejecting *Evidential Equivalence* comes at a very serious cost, and it perhaps raises many more questions than it answers. This should not be too surprising, given that *Evidential Equivalence* is more or less built into the foundations of Bayesian epistemology. In closing, I will briefly state three initial worries for the denier of *Evidential Equivalence*.

First, there are nearby cases where it is not at all clear what the denier of *Evidential Equivalence* should say. For example, what should $\text{Cr}(H_A \mid E_1 \vee E_{2,A})$ be? Should it be 0 or 1/2? Furthermore, what should my credence in H_A be if I first updated on E_1 and then update on $E_{2,A}$? Would this change if I first updated on $E_{2,A}$ and then updated on E_1 ? Is the commutativity of conditionalization violated? There is also the general worry that rational agents shouldn't have to wait to be 'told' the second piece of information, since they can immediately infer the second piece of information from the first.

Second, one will need a hyperintensional account of the objects of credence. On one standard view, the objects of credence are something like sets of possibilities, where the possibilities are either epistemic or metaphysical possibilities.¹⁶ Since E_1 and $E_{2,A}$ describe the same set of possibilities, any account like this will entail proposition 3. One natural thought is that the objects of credences should be something like sentences, since there are clearly hyperintensional distinctions between necessarily equivalent sentences. If this sort of account is adopted, one must decide whether the sentences should come from some natural language or some highly idealized language. Another account is the one defended by Braun (2016), in which the objects of credence are Russellian propositions which are composed from the individuals and properties that one's belief is *about*. This will have the consequence that E_1 and $E_{2,A}$ are distinct objects of credence since the first is partially about A and the second is not. Lastly, one might try to adopt the formal framework in Fine's (2017) truthmaker semantics by letting the objects of credence be sets of 'states' as opposed to sets of worlds. Fine's framework is explicitly hyperintensional, and it accommodates the phenomenon of aboutness particularly well (p. 569-71).

Third, *Evidential Equivalence* can be derived from central principles about conditional probability.¹⁷ So, if *Evidential Equivalence* is denied, then certain central principles about conditional probability will also have to go. The proof uses the following axiom of conditional probability, which is present in the axiomatizations of primitive conditional probability found in Popper (1955) and Rényi (1970):

$$\text{Multiplicative Axiom (MA): } \text{Cr}(A \wedge B \mid C) = \text{Cr}(A \mid B \wedge C)\text{Cr}(B \mid C)$$

Suppose that one is certain that E_1 is true iff E_2 is true, as per *Evidential Equivalence*. Then, one should set $\text{Cr}(E_1 \mid E_2) = 1 = \text{Cr}(E_2 \mid E_1)$. By MA, $\text{Cr}(H \wedge E_2 \mid E_1) = \text{Cr}(H \mid E_2 \wedge E_1)\text{Cr}(E_2 \mid E_1) = \text{Cr}(H \mid E_2 \wedge E_1)$. Similarly, by MA, $\text{Cr}(H \wedge E_1 \mid E_2) = \text{Cr}(H \mid E_1 \wedge E_2)\text{Cr}(E_1 \mid E_2) = \text{Cr}(H \mid E_1 \wedge E_2)$. So, since $\text{Cr}(H \mid E_2 \wedge E_1) = \text{Cr}(H \mid E_1 \wedge E_2)$, it follows that $\text{Cr}(H \wedge E_2 \mid E_1) = \text{Cr}(H \wedge E_1 \mid E_2)$.

¹⁶ For a defense of the view that the objects of credence are sets of epistemic possibilities, see Chalmers (2011). For a traditional defense of the view that the objects of credence are (centered) possible worlds, see Lewis (1986).

¹⁷ Thanks to an anonymous referee for the following argument.

Next, note that $\text{Cr}(H \wedge E_2 | E_1) = \text{Cr}(H | E_1) + \text{Cr}(E_2 | E_1) - \text{Cr}(H \vee E_2 | E_1)$, and since $\text{Cr}(E_2 | E_1) = 1 = \text{Cr}(H \vee E_2 | E_1)$, we have that $\text{Cr}(H \wedge E_2 | E_1) = \text{Cr}(H | E_1)$. Exactly symmetric considerations imply that $\text{Cr}(H \wedge E_1 | E_2) = \text{Cr}(H | E_2)$. So, we have $\text{Cr}(H | E_1) = \text{Cr}(H \wedge E_2 | E_1) = \text{Cr}(H \wedge E_1 | E_2) = \text{Cr}(H | E_2)$, as desired. Since each of the steps in this proof proceeded through central, uncontroversial principles governing conditional probability, rejecting *Evidential Equivalence* comes at a very steep cost.

All this being said, it seems like *every* way out of our inconsistent triad comes at a very steep cost. In order to deny proposition 1, you would need to say that learning that *almost everybody flipped tails* is no evidence at all that *you* flipped tails. Moreover, you would need to choose to be more inaccurate, to expose yourself to dutch books, to ‘wait’ until an announcer tells you the exact number of head-flippers, and to refuse to defer to ideally rational epistemic experts. In order to deny proposition 2, you would need to accept the claim that evidence entirely about the outcomes of *other* coin tosses is somehow relevant to the outcome of your *own* coin toss, *holding fixed* that every coin toss is independent of every other one! Moreover, you would need to either find some disanalogy between Finite Coins* and Past Coins or have a bizarre account of ‘inadmissible’ evidence in Past Coins. Neither of these two options is particularly attractive.

Whatever one ultimately decides to say about this paradox, I hope to have pointed out an interesting tension in the foundations of Bayesian Epistemology. While there are principled reasons why we should uphold *Evidential Equivalence*, there are also principled reasons for thinking that certain central concepts in Bayesian epistemology, namely *independence* and *inadmissibility*, are intuitively hyperintensional. In some way or other, this tension needs to be addressed.¹⁸

References

- Braun, David 2016. The Objects of Belief and Credence. *Mind* 125, no. 498: 469–497.
- Chalmers, David J. 2011. Frege’s Puzzle and the Objects of Credence. *Mind* 120, no. 479: 587–635.
- Dorr, Cian. 2010. The Eternal Coin: A Puzzle About Self-Locating Conditional Credence. *Philosophical Perspectives* 24, no. 1: 189–205.

¹⁸ Thanks to David Balcarras, Zach Barnett, Kenny Easwaran, Caspar Hare, Agustín Rayo, Haley Schilling, Miriam Schoenfield, Jack Spencer, Roger White, Stephen Yablo, and two anonymous referees for helpful feedback and discussion.

- Easwaran, Kenny. 2008. *The Foundations of Conditional Probability*. PhD thesis, University of California, Berkeley.
- 2013. Why Countable Additivity? *Thought: A Journal of Philosophy* 2: 53–61.
- 2014. Regularity and Hyperreal Credences. *Philosophical Review* 123, no. 1: 1–41.
- Fine, Kit. 2017. Truthmaker Semantics. In *A Companion to the Philosophy of Language*, edited by Bob Hale, Crispin Wright, and Alexander Miller, 556–577. Blackwell, New York.
- Hájek, Alan. 1996. ‘Mises Redux’ — Redux: Fifteen Arguments Against Finite Frequentism. *Erkenntnis* 45, no. 2–3: 209–27.
- 2003. What Conditional Probability Could Not Be. *Synthese* 137, no. 3: 273–323.
- 2009. Fifteen Arguments Against Hypothetical Frequentism. *Erkenntnis* 70, no. 2: 211–235.
- Hawke, Peter. 2018. Theories of Aboutness. *Australasian Journal of Philosophy* 96, no. 4: 697–723.
- Kadane, J. G., Schervish, M. J., and Seidenfeld, T. 1986. Statistical Implications of Finitely Additive Probability. In *Bayesian Inference and Decision Techniques*, edited by P. Goel and A. Zellner, 59–76. Amsterdam: Elsevier.
- Krämer, Stephan. 2017. A Hyperintensional Criterion of Irrelevance. *Synthese* 194, no. 8: 2917–2930.
- Leitgeb, H. and Pettigrew, R. 2010. An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science* 77: 201–35.
- Lewis, David. 1980. A Subjectivist’s Guide to Objective Chance. In *Studies in Inductive Logic and Probability*, edited by Richard C. Jeffrey, 83–132. University of California Press.
- 1986. *On the Plurality of Worlds*. Wiley-Blackwell.

Myrvold, Wayne C. 2015. You Can't Always Get What You Want: Some Considerations Regarding Conditional Probabilities. *Erkenntnis* 80, no. 3: 573–603.

Popper, Karl. 1955. Two Autonomous Axiom Systems for the Calculus of Probabilities. *British Journal for the Philosophy of Science* 6, no. 21: 51–57.

Rényi, Alfred. 1970. *Foundations of Probability*. Holden-Day.

Yablo, Stephen. 2014. *Aboutness*. Princeton University Press.