# Defeating Objections to Bayesianism by Adopting a Proximal Facts Approach

*Calum Miller*
*Uehiro Centre for Practical Ethics, University of Oxford*

## Introduction

One major line of attack against probabilistic approaches to the philosophy of science has been to argue that certain results of theirs are in conflict with intuitive notions of confirmation. Thus for example, some have suggested not only that the Hempelian raven paradox[1] counts against standard, prep-robabilistic notions of scientific confirmation but also that it demonstrates a problem with approaches based on confirmation theory: since P(nonblack object being a nonraven | all ravens are black) is 1, it follows from Bayes's theorem that the observation of a nonblack nonraven constitutes evidence that all ravens are black.[2] Those who find the raven paradox persuasive, and who retain their intuition that such an observation does not even slightly confirm the black raven thesis, ought to find this a compelling argument against Bayesianism, for the probabilistic account contradicts the ostensible common-sense intuition.

Others see this as a strength of Bayesianism—that Bayesianism accepts the otherwise plausible equivalence condition[3] yet also accounts for the fact that we do not hold such observations to significantly confirm the black raven thesis. The reason for this is that the probability of a nonblack object being a nonraven given that *not* all ravens are black is trivially close to 1, even though it is not 1. This means that the observation—a nonblack nonraven—is to be expected with a high degree of probability *regardless* of whether all ravens are black. So the increase in the epistemic probability of the black raven thesis is negligible.

---

[1] Carl G. Hempel, "Studies in the Logic of Confirmation II," *Mind* 54, no. 214 (1945): 97–121.

[2] This assumes, of course, that P(nonblack object being a nonraven | ~[all ravens are black]) < 1, which is trivially true.

[3] The equivalence condition says that if X is evidence for Y, then X is evidence for anything logically equivalent to Y. Hence "all ravens are black" is logically equivalent to "all nonblack things are nonravens."

## Inverse Gambler's Fallacies and Bayesianism

Whatever one's position on the raven paradox, it is clear that these kinds of intuitive conflicts may in theory constitute a major group of objections to Bayesianism. Let us turn to the difficulty I present in this essay. Take the datum "tables exist"—a relatively noncontroversial proposition. And let our hypothesis M be that there exists a multiverse—an enormous number of disjoint space-times, or universes.[4] According to the Bayesian account, our observation of tables existing would thus seem to confirm the existence of a multiverse. This is because the probability of tables existing is greater given the existence of a multiverse than the probability of tables existing in the absence of a multiverse. Formally, where $T =_{df}$ tables exist, $P(T|M) > P(T|{\sim}M)$, and thus by Bayes's theorem, $P(M|T) > P(M)$. The former inequality can be justified simply by appealing to the fact that the more space-times there are, the higher the likelihood that tables would exist in at least one of them and thus the higher the probability that the proposition "tables exist" is true. The latter inequality just says that the probability of the proposition "a multiverse exists" is raised by the truth of the proposition "tables exist." So on the standard Bayesian account of evidence, T is evidence for M.

This can, it seems, be applied to any existential proposition. The aforementioned reasoning would suggest that any existential proposition, if true, would provide evidence for the multiverse hypothesis. Thus "plants exist" and "sandwiches exist" would also constitute evidence for M. The existence of absolutely anything would seem to confirm M, and thus we would seem to have overwhelming, perhaps infinite, evidence for M. Moreover, this is not trivially confirmatory evidence, as in the raven paradox. Rather, the existence of even a pen would seem to strongly confirm M, since we might suppose P(a pen exists|~M) to be reasonably low, while P(a pen exists|M) would be moderately high.

Worse still, objects' existence would confirm M even more insofar as the object is specified in a more detailed way. For example, let $Rx =_{df}$ x is a red pen in a pot, along with several other pens, on a kitchen counter, in a terraced house, in a city whose name begins with "L," in a continent whose name begins with "E," on a planet whose name begins with "E," in the Milky Way.[5] Then $P((\exists x)Rx|{\sim}M)$ would seem to be extraordinarily low, whereas $P((\exists x)Rx|M)$ would be moderately high, provided M was a sufficiently extravagant

[4] The precise nature of the multiverse hypothesis under consideration is irrelevant.

[5] Evidently I wrote this sitting at my desk in London; more imaginative philosophers will be able to cite more interesting examples.

multiverse hypothesis. This would imply that (∃x)Rx acts not only as one out of many pieces of evidence for M but counts as *overwhelmingly strong* evidence for M. It seems that any mundane object we come across could be described in such a detailed fashion, and so could each constitute extraordinary evidence for M.

Finally, it is not clear that the evidential force of each object's existence would be greatly limited by the evidential force of other objects' existence. The Bayesian approach typically shows that the evidential force of some observations is limited by the evidential force of previous similar observations. For example, if one already knows the results of a particular experiment after a thousand repetitions, the results of the 1,001st repetition will not dramatically alter the relevant hypotheses' epistemic probability. In Bayesian terms, this is because the results of the first thousand experiments ought to lead us to expect the same result no matter whether the hypothesis we are testing is true or not—and so there will be very little confirmation by repeating the same experiment. Let us call this phenomenon *evidential limitation*.[6]

But there does not seem to be a huge amount of evidential limitation here.[7] For even though (∃x)Rx entails the existence of some other things—other pens in the pot, for example, one can easily find the existence of an enormous number of objects, largely independent, each of which confirm M immeasurably. In addition to (∃x)Rx, then, one might also offer D in support of M, where $D =_{df}$ there exists a precisely shaped dent (specified in detail) in the biggest planet in a universe. P(D) does not seem to depend much on P((∃x)Rx), and so would seem to give independent support to M—and again, this would seem to be overwhelmingly strong support.

Some might perceive the primary difficulty here to be the seeming commitment, from such trivial data, to an unfavorable metaphysical position—namely, the multiverse. Even if the prior probability of a multiverse is minute, surely the evidence previously described is sufficient to

[6] This phenomenon arises because the evidence is not *independent*—that is, the probability of the latter data is affected by the knowledge of the first piece of data.

[7] There is some evidential limitation. For physicists have shown—by apparent consensus—that the laws of physics had to be very precise in order for any complex matter to exist at all. Thus the laws of physics are said to be "fine-tuned." Insofar as the existence of material objects requires such laws, the existence of some material object will significantly raise the probability of other physical objects existing. Nevertheless, since the latter physical objects can be specified in arbitrarily more detailed ways, the probability of their existence is still low even given the existence of other material objects.

overcome it. And this seems troubling for those who are averse to a multiverse, for whatever reason.

But one solution might be to consider this a *reductio ad absurdum* against the multiverse: if sound probabilistic reasoning leads us to consider these phenomena as extremely strong evidence for the multiverse, perhaps the fact that we do not *really* consider a multiverse to probably exist only serves to demonstrate that the prior probability of a multiverse *must* be infinitesimally low. This proposed solution would account for our intuitive conviction that we are not really compelled to accept the multiverse on such trivial grounds.

I have some sympathies with this view, and it is plausible that quantitative parsimony is indeed a theoretical virtue.[8] Thus a multiverse has lower prior probability insofar as it becomes more extravagant.[9] But this kind of Moorean shift, in this case, is entirely inadequate. For one thing, many will consider such a move to be extremely *ad hoc* and will feel that staunch Bayesians who do not have independent reason to appraise P(M) as insuperably low *must* accept M.

But a more worrying problem arises for Bayesians: even those who have no problem with accepting M are faced with the difficulty that the confirmation of M provided by these existential claims is at odds with intuitive notions of confirmation. Indeed, such confirmation seems to be a clear instance of a well-known probabilistic fallacy: the inverse gambler's fallacy.

The paradigmatic case of this fallacy involves a gambler who walks into a casino and immediately sees that a 6 has been rolled on a die. This, he thinks, suggests that the die has been rolled a large number of times. After all, P(a 6 is rolled on the die | the die has been rolled many times) is far greater than P(a 6 is rolled on the die | the die has been rolled only once). Indeed, the critique of Bayesianism currently under consideration may adduce this as a further example: according to Bayesianism, this datum would significantly confirm the many-rolls hypothesis over the single-roll hypothesis (one can amplify the strength of such ostensible evidence by using an even more improbable example: a result of 35 on a roulette wheel, for example). But clearly such an inference is absurd: we realize that the gambler was always bound to see *something* rolled on the die, and it just happened to be a 6 in this case. It is hardly clear that his seeing a 6 on the first roll he sees confirms the many-rolls hypothesis more than any other number would—but then we are

---

[8] See Daniel Nolan, "Quantitative Parsimony," *British Journal for the Philosophy of Science* 4, no. 3 (1997): 329–43, and Michael Huemer, "When Is Parsimony a Virtue?," *Philosophical Quarterly* 59, no. 235 (2009): 216–36.

[9] At least insofar as it is posited as a brute fact: if it is a consequence of a simpler or more well-evidenced mechanism, then a multiverse may not become monotonically more implausible as its constituents multiply.

left with the possibility that *any* result would confirm many rolls or that *none* of the results confirms it. The first of these options seems clearly wrong, since he would see *some* result whether there had been many rolls or just the one. But then the only option is that no particular result confirms the many-rolls hypothesis, and this is what we generally take to be true.

The problem, then, is that the Bayesian approach seems to contradict this account and that the multiverse example is another prime instance of this fallacy. For it seems that we were bound to see that *some* things exist so long as only one universe exists (and has observers) and, intuitively, we should be just as likely to observe tables and pens if only our universe existed, as if many universes existed. If the other universes are not observable to us, then they should have no impact on how likely it is that we see particular objects existing. But if we use these existential claims as data in our Bayesian schema, then we seem compelled to accept that they confirm the multiverse hypothesis. The central problem can be summarized thus: the Bayesian approach seems to advocate confirmation in these inverse gambler's fallacies, when really there is none.

## A Possible Solution

All Bayesians are therefore liable to this objection and must come up with some solution if they are to maintain that Bayesianism is consistent with standard epistemological praxis. One more promising solution might be suggested by reflecting on the standard inverse gambler's fallacy: we realize, in the case of the gambler, that the fact *can* be construed so as to avoid any confirmation. We noted that we would expect the gambler to see *something* as he walks into the casino and sees his first die roll, and in this case it happened to be a 6. So although it is the case that P(a 6 is rolled on a die | the die has been rolled many times) > P(a 6 is rolled on a die | the die has been rolled only once), it is not the case that P(the gambler observes a 6 | the die has been rolled many times) > P(the gambler observes a 6 | the die has been rolled only once), provided the relevant background information is also included in the conditional (i.e., that he has entered a casino and that this is the first roll of the die he has witnessed). This is because *what the gambler observes* is not likely to depend at all on how many times the die has been rolled previously, even if the *existence* of some roll of a 6 does depend on it.

So it can be seen that how data are construed can make a huge difference to the degree to which they confirm a certain hypothesis. Perhaps, then, the Bayesian will want to restrict Bayesian propositions in some way. One way this could be done is to restrict propositions to observational propositions. Thus "I observe a pen" might be an appropriate proposition to appraise, whereas “a pen exists” is not. This could potentially go a long way toward alleviating the ostensible commitment to inverse gambler’s fallacies, but it faces some forceful questions. First, what type of observational statements are most appropriate? Second, what rationale can one give for using these observational data and construing them this way rather than some other data or a different construal? Third, are we to limit *all* Bayesian propositions (including, for example, the hypotheses that we are supposed to be appraising) to these kinds of observational propositions? And finally, if not, by what criteria can we judge when nonobservational propositions are appropriate, and is it possible to find non–*ad hoc* criteria that do justice to the scientific method? If this kind of solution is utilized, it ought to be able to answer these questions persuasively.

I submit that the optimal solution lies in restricting the data used in confirmation to what I will call “proximal facts.” This, rather than simply providing some *ad hoc* solution to the aforementioned paradoxes, is the most natural and appropriate way of utilizing Bayesian confirmation and can be demonstrated to be so by the classic problem of uncertain learning.

## Uncertain Learning and Proximal Facts

Suppose there is a murder case, and you, the detective, have been called to the crime scene to view any evidence that might help in your investigation. On reaching the crime scene, you come across what you immediately recognize as one of Mr. Wood’s kitchen knives—and indeed, after a brief assessment, you are confident (i.e., think it probable) that it is his knife (call this fact—that Mr. Wood’s kitchen knife is found at the scene—K). Since Mr. Wood is already one of the primary suspects, you are inclined to think that this provides significant confirmation of his being the culprit (W). Being a good Bayesian, you form a probabilistic assessment and conclude that $P(K|W) >> P(K|{\sim}W)$ and that $P(W|K)$ is now very high. You leave the knife at the crime scene and return to the station to contact Mr. Wood.

As it happens, Mr. Wood is currently in some distant country and so you are limited to a video interrogation with police at the other end. His quick escape puzzles you, but even more so when he shows you his kitchen knife. When the police with him verify that it is, indeed, his (P), you are astounded, for you saw the same knife a very short time ago, far too short for it to have travelled abroad since. You rush to find a way of reconciling these, coming up with one implausible hypothesis after another—that he quickly boarded an airplane faster than ever before, that he somehow teleported abroad, that some deity intervened to bring this about, and so on. After all, you *know* that his knife was with you at the crime scene just recently, and so you must come up with some hypothesis that takes account of this and includes it among its explananda. All these hypotheses seem *a priori* unconvincing to you, but you must search for the most probable among them in light of the evidence.

Of course, in such a situation, where all the remaining explanations that account for the data are so staggeringly improbable, the prudent detective will recheck his facts. Despite his initial (and reasonable) assessment that the knife belonged to Mr. Wood, most would agree that the more probable situation, *in light of these other facts*, is that the knife did not really belong to Mr. Wood at all. Most of us recognize that, rather than adopt a phenomenally improbable hypothesis just to account for what we consider to be probably true facts, we should more readily give up one or more of those facts, even if they seemed to us more probable than not before considering their implications (and even if we thought we *knew* them). And *ceteris paribus*, we should be more inclined to give up those facts that seemed only slightly more probable than not.

In probabilistic terms, the detective has found that, despite all remaining hypotheses (call these hypotheses $A_1 \ldots A_n$, whose disjunction might be called A) having inconceivably low prior probability, nevertheless $P(A|K \& P)$ is moderate. But rather than think that the most probable of these *ad hoc* hypotheses are true, or even that a disjunction of some of them is true, we seem more compelled to give up K and to remove it from our conditional. This is an instance of the "problem of uncertain learning,"[10] and a good solution ought to make sense of this situation as well as other instances.

But oughtn't K be replaced by something? After all, we cannot simply ignore the knowledge we have, so it seems reasonable that we include it somehow. But in light of P, K seems to become much more improbable. Since we were originally less certain of K than we are of the falsehood of $A_1 \ldots A_n$, we have given it up—and with full awareness that our primary circumstantial data might not be as it seems, we now look for more certain and well-founded data with which we can confirm or disconfirm our hypotheses. Thus we might instead use the datum "a kitchen knife resembling Mr. Wood's was found at the crime scene" (K'). This is surely much more well-founded than K and does not entail (when conjoined with P) that a disjunction of absurd, *ad hoc* hypotheses is true. Indeed, one might much more easily find some priorly plausible hypothesis H, accounting for K' & P, such that

[10] For a helpful discussion of uncertain learning in the context of uncertain auxiliary hypotheses, see Michael Strevens, "Notes on Bayesian Confirmation Theory" (June 2017), http://www.nyu.edu/classes/strevens/BCT/BCT.pdf.

$\sim[P(H) << 1]$[11] and $P(H|K' \& P) > 0.5$.[12] One such hypothesis is that the other primary suspect, Mrs. Walker, who is widely known to resent Mr. Wood, bought a kitchen knife resembling Mr. Wood's and used it to commit the crime, thus framing Mr. Wood at the same time. Such a hypothesis would be rendered extremely unlikely (or impossible) conditioned on K, but on K' it seems most reasonable.

K can still be included as a possibility in this circumstance and can be worked into the probability assessments. In this case, K would work as an uncertain auxiliary hypothesis in the same way that W is an uncertain hypothesis, rather than as a datum taken as certainly true. Thus in our example, $P(K|K')$ would be moderately high—accounting for our initial judgment—but $P(K|K' \& P)$ would be very low indeed.

To show how uncertain data such as K can still be included, and how using a much more certain datum instead (K') can give us a more realistic probability assessment, let us consider how such a probability assessment might work. We have said that P(A) is negligible, but $P(A|K \& P) = 1$. If K and P are to be taken as certainly true, then we have no choice but to accept that A is true—that is, that some absurd, inconceivably implausible hypothesis (though we know not which) is, given the data, correct. But if K' is used instead, we can calculate $P(A|K' \& P)$, accounting for the uncertain auxiliary hypothesis K thus:

$$P(A|K' \& P) = P(A|K \& K' \& P)\cdot P(K|K' \& P) + P(A|{\sim}K \& K' \& P)\cdot P({\sim}K|K \& P)$$

Here, even though $P(K|K')$ is high, we are asked to calculate the probability of K not just given K' but also given P. And when P is considered, we have much less confidence in K. Thus $P(K|K' \& P)$ is inconceivably low (and $P({\sim}K|K' \& P)$ is nearly 1). This will effectively make the whole first conjunct negligible, leaving a close approximation:

$$P(A|K' \& P) \approx P(A|{\sim}K \& K' \& P)$$

And it is far from clear that $P(A|{\sim}K \& K' \& P)$ is anywhere near 1. Indeed, if P(A) is negligible and we have at least one other plausible hypothesis under ~A that is compatible with ~K & K' & P (a much easier condition to fulfill than compatibility with K & P, since it allows for a similar-looking knife to have been found at the crime scene—a relatively plausible proposition), then it seems that, on this account, $P(A|K' \& P)$ and $P(A|{\sim}K \& K' \& P)$ remain

[11] That is, such that P(H) is not too much less than 1 (i.e., is somewhat plausible).

[12] For example, that the probability of H given the data is greater than 0.5.

incredibly low. So we arrive at the intuitive result: $P(A|K' \& P)$ is low. This, unlike our initial account, fits far better with our intuitive understanding of theory appraisal and shows how K can nevertheless be fitted into the schema as one possible (and priorly probable) hypothesis but also how the prior plausibility of K can be overcome such that we can reject it if necessary. Thus we can avoid concluding that absurd, counterintuitive hypotheses are correct.

So we have an account here that accords with our intuitions and that can be generalized to all other cases of uncertain learning. And it suggests that we should conditionalize on the most certain facts possible. This means that facts like K should be disregarded as *data* in favor of facts like K'. This approach of looking for the most certain facts is what I call the proximal facts approach to Bayesianism.

What will such an approach involve? It seems that even highly probable data such as K' will prove inadequate for a comprehensive Bayesian analysis, though they may be useful in practice. This can be illustrated by modifying our original example. Suppose that, after modifying your datum to K' and having become satisfied with this more proximal fact, you then come across some more information—when discussing your case with your team, they have no recollection of any knife being found at the crime scene. Moreover, when you next return to the scene, you find no trace of any knife, despite high, trustworthy security. On returning home, you find what appear to be hallucinogens sitting on your kitchen counter beside your early morning coffee cup. Call these data H.

Again, we have a similar situation to before. In light of H, K' itself seems very implausible, and intuitively we should no longer accept K'. Rather than trying to come up with a new set of implausible hypotheses that account for K' & H, then, we recognize that it may seem better to abandon K' in spite of the fact that P(K'|your initial observations) is high. K', though a far more certain fact than K, is now improbable itself in light of H. So to provide even more certainty, you appeal to K'', a fact that is surely as certain as any other: "I have a percept (whether veridical or not) of a knife that seems to me to resemble Mr. Wood's." Rather than having to face another disjunction of wildly improbable explanatory hypotheses (call these $A_2$) so as to allow K' & H to feature in our probabilistic antecedent (which would give $P(A_2|K' \& H) = 1$), we are faced with a new analysis:

$$P(A_2|K'' \& H) = P(A_2|K' \& K'' \& H)\cdot P(K'|K'' \& H) + P(A_2|{\sim}K' \& K'' \& H)\cdot P({\sim}K'|K'' \& H)$$

As with the previous case, $P(K'|K'' \& H)$ will now be seen to be negligible despite $P(K'|K'')$ being relatively high. Since this will make $P({\sim}K'|K'' \& H)$ roughly equal to 1, we are again left with:

$$P(A_2 | K'' \& H) \approx P(A_2 | {\sim}K' \& K'' \& H)$$

And once again, this is nowhere near 1 because we no longer have to account for K' & H. Again, this delivers the right result. Having only to account for K" & H, we have a much better hypothesis not included in $A_2$—that you hallucinated the knife at the crime scene. And so we can see, once again, that including only the most proximal facts—that is, those most certain to us—provides the most complete and compelling analysis. This latter example demonstrates that even data such as "I observe a knife" are inadequate (insofar as "observing" is understood as entailing veridicality). In order to include the possibility that even our basic observations are mistaken interpretations, we must appeal to the most immediate facts—that is, our ostensible percepts of certain phenomena. Thus the proximal facts approach to Bayesianism should be using this basic perceptual information as its data, with nonproximal facts being used as data only for practical and expediential purposes.

## Finalizing a Defense of Bayesianism

One can now see how this natural, powerful, and comprehensive approach to Bayesianism may provide a compelling solution to the original objection. In particular, our first two questions seem now to have decisive answers. In response to the first, the type of observational statements that are most appropriate are those that constitute the most fundamental, certain aspects of our experience—namely, the data describing our percepts, with no implied appraisal of the veracity of these percepts. And in response to the second, the rationale is that using these data allows us to account for cases of uncertain learning. Doing so also allows us to take all the explanatory possibilities into account when assessing the probability of different hypotheses rather than limiting our pool of possibilities by assuming the certainty of given propositions that are really less than certain and that are potentially defeasible.

Moreover, we can now begin to formulate an answer to the final two questions. When reflecting on the inverse gambler's fallacies, we noted that we might wish to limit our propositions to those pertaining to observation. But if this is the case, should we make *all* the propositions included in our Bayesian framework these kinds of proximal facts? Including those hypotheses that we are supposed to be using observational evidence to confirm or disconfirm? If so, what is the use of Bayesianism? All our hypotheses would be claims about what we observe—but we already know what we observe, and the whole point of scientific investigation (at least, for the realist) is to use observed data to help appraise propositions regarding what we do not observe! But if we shouldn't restrict *all* Bayesian propositions in such a way, then how do we rationalize our decision to allow nonproximal propositions and nonobservational data anywhere in our schema?

The answers to these questions come from a consideration of the purpose of Bayesianism and its intended use in science. A brief overview of how it is alleged to work is therefore necessary. According to Bayesians, science is characterized by a sharp distinction between two kinds of propositions to be appraised: those recording observed phenomena and those not recording observed phenomena. The idea, for scientific realists at least, is that observed data somehow determine, or contribute to, our appraisal of the unobserved. In some cases, the unobserved will be propositions regarding what will happen in the future; in other cases, the propositions will pertain to the behavior of entities not directly observed. The classic problem of induction (as well as a host of other riddles) relates to these and concerns how one justifies the generalization from observed data (say, the color of emeralds or the behavior of atoms) to unobserved situations (such as the color of emeralds in the future or the behavior of atoms elsewhere in the universe). In addition to this, there are a wide range of theories purporting to explain these data, which are generally not held to be directly observed, or entailed by observed data, but justified by some other method. Thus the truth of Einstein's general theory of relativity is not directly observed but is held to give a good explanation of observed data and is also considered a relatively parsimonious explanation. The problem of how to decide between these "unobservable" theories is the subject of theory appraisal and is typified by the problem of underdetermination.

How observations are supposed to support a theory or not has been the subject of enormous controversy for centuries, with all the proposed solutions being attempts to find a rational justification for holding theories to be true, despite arguments from the data to the theories generally being deductively invalid. Bayesianism attempts to demonstrate the rationality of such moves by appealing to the probability calculus. Here, so long as certain axioms are granted (particularly the definition of conditional probability), it can be demonstrated that for any h and e, $P(h \mid e) = P(h) \times P(e \mid h) / P(e)$, where h represents the hypothesis and e represents the observational evidence. This equation, Bayes's theorem, demonstrates that the epistemic probability of a hypothesis given the observational evidence is higher than the probability of the hypothesis before[13] considering that evidence, if the evidence is more to

[13] "Before" is here used in a logical, rather than temporal, sense and is not intended to connote any allegiance to predictivism—the thesis that novel predictions confirm a theory more than previously known observations that are similarly accommodated or "predicted" in retrospect by the theory. It is a matter of some debate

be expected given the hypothesis than otherwise. This makes sense of the notion that observations "confirm" a hypothesis if they are predicted by that hypothesis and more so insofar as they are unexpected otherwise. The hope, then, is that one will eventually come up with an epistemic probability $P(h|k)$, where k includes all observations so far made, in order to judge how probably true the hypothesis is.

It seems clear, then, that there is no comparable reason to limit *these* propositions to observational propositions. For the whole point is that we are considering the epistemic probability of the hypothesis, given *the data*. The reason we require proximal facts at all is that we need certainty (or as near as possible) for the data that form the probabilistic antecedent of our conditional probability in order to attain the most accurate appraisal of a certain hypothesis. But since we are not including these uncertain hypotheses in our antecedent knowledge, we have no need of their certainty—the whole point is that we are trying to assess their probability, without assuming that they are known for certain. So it is the need for certainty (or as close as we can get) that is the criterion by which we require some propositions in the Bayesian analysis to be of a particular kind, and this provides the answer to our final question.[14]

## Implications

We are now in a position to demonstrate why this most plausible reading of Bayesianism is not committed to endorsing the fallacies described earlier. Let us take the example of the gambler first. Instead of the gambler walking into the casino with his background knowledge and using the datum "a 6 has been rolled on a die," which *would* give confirmation of the many-rolls hypothesis, the gambler is rationally compelled to use the more proximal datum "*it seems that* I observe that a die has been rolled to give a 6." This, unlike the previous

whether predictivism is true, though I am heavily inclined to think not. See Richard Swinburne, *Epistemic Justification* (Oxford: Oxford University Press, 2001), 221ff.

[14] A clarification is required here. I am not saying that only the most proximal facts can feature in *any* probabilistic antecedent (i.e., in the conditional). I am only saying that when we are trying to determine on what we should conditionalize when trying to determine our credence in some proposition, we should use such facts—that is, if we want to know what our credence in F should be and E represents all our evidence as proximal facts, we should use the probability $P(F|E)$. That is, of course, compatible with conditionalizing on other propositions as part of the calculation. Indeed, this is necessary for Bayes's theorem in the first place, since part of the calculation includes $P(E|H)$, where H—the hypothesis—is far from certain!

datum, does not seem to give any confirmation to the many-rolls hypothesis over the single-roll hypothesis, and no fallacy is committed.

In the case of the multiverse, we are in a similar position. Instead of using the more distal, uncertain facts, like "tables exist" and so on, the rational observer ought to instead use, "it seems that there is something resembling a table in my visual field" or something similar. Again, unlike the previously construed datum, this does not seem to give any confirmation to the multiverse hypothesis, since the multiverse hypothesis does not seem to affect what we observe at all.

We have, therefore, a resolution of the problems set out at the beginning of the essay, and we have a well-motivated framework for what we should include in our conditional probabilities. But there is one final implication I would like to draw out related to our thinking about cosmic fine-tuning.

Although, given my embodied existence,[15] it is just as likely that I would observe the existence of tables (or any other given material object or detail) given a multiverse than given only one universe, we might reasonably ask whether my embodied existence itself is evidence for the multiverse. And it appears that it is. For P(I observe that I exist|multiverse) appears to be significantly larger than (I observe that I exist|universe). We cannot use the same reasoning as in the other cases, for it is *not* inevitable that I observed *something*—I might not have existed at all! And if the multiverse makes it more likely that I observe my existence by making it more likely that I exist in the first place, then it seems as though my existence is evidence for the multiverse.

This is not as counterintuitive as my initial examples. Indeed, since the discovery of cosmic fine-tuning, many physicists and philosophers have been keen to posit a multiverse as an explanation of why the laws of the universe are thus. So this need not be a problem for Bayesians. Indeed, it appears as though whichever precise construal of the datum is used in the conditional, the multiverse makes embodied life more likely than otherwise. Our intuitions and the Bayesian result coincide: the existence of embodied life, given our knowledge of fine-tuning, is evidence for a multiverse. But given this fine-tuning and the existence of embodied life, observations of other material objects or details is not so surprising: it seems as though we can observe just as much whether there is one universe or many, since we can only observe the universe we are part of, even if there are many. This is analogous to the

[15] The assumption of embodied existence is a necessary simplification here. If it were possible for me to exist unembodied (and so not requiring fine-tuned laws of physics), then a multiverse would perhaps make it much more likely that I observe material objects than if there were only one universe. But here I will assume that I am essentially embodied.

gambler: he will witness the same thing whether or not there have been many rolls of the die before the first roll he witnesses. But imagine we change the gambler case to make it relevantly analogous. Given the nature of the case, this will be somewhat contrived. But suppose that the gambler witnesses not a 6-sided die roll but a lottery with ten million possible outcomes. And suppose his entry into the casino was dependent on one particular outcome of this lottery. To be sure, if he does enter the casino, he will not be surprised to witness this outcome. But he should nevertheless be surprised that he entered the casino in the first place if only one lottery took place. But if many lotteries took place, it is much more likely that he would enter the casino at some point. So in this case, it does not seem as though he has committed the inverse gambler's fallacy—his inference that there were probably many lotteries is a reasonable one. Likewise, our existence is evidence for a multiverse.

How strong the evidence is depends on other considerations. In general, the strength of a piece of evidence E for a hypothesis H is limited by P(E)—the probability of the evidence in general. This, in turn, is a function of the probability of the evidence given each member of a partition (i.e., a set of mutually exhaustive and exclusive propositions) of possible explanations for that evidence. These probabilities are to be weighted depending on the prior probability of those alternative explanations. Thus:

$$P(E) = \sum P(E \mid H_i) \times P(H_i).$$

The lower P(E) is, the stronger evidence E is for H. So if there is another possible explanation $H_i$ of E that has a moderate prior probability and that leads us to expect E—that is, $P(E \mid H_i)$ and $P(H_i)$ are both moderate—this will limit the strength of the evidence. In the case of fine-tuning and our embodied existence, some philosophers and scientists suppose that they have alternative explanations, whether theism, some deeper scientific explanation, or something else. And in the case of fine-tuning arguments for theism, the fact that our existence is evidence for the multiverse does not undermine such arguments. For just as our existence is evidence for the multiverse, in the same way it is evidence for theism (given certain assumptions that I do not have space to defend here). And again, the strength of this evidence for theism will be limited by the prior plausibility of alternative hypotheses—like the multiverse itself. But note that the evidential force is limited only by the *prior* plausibility of such alternative hypotheses—that is by their intrinsic plausibility or by *independent* evidence for them. Since our existence is not a piece of *independent* evidence but the piece of evidence to be explained, the fact that it also supports a multiverse theory will not limit the strength of the fine-tuning argument for theism. And as I have argued, since the existence of other material objects does not support the multiverse theory once the initial datum of fine-tuning has been taken into account, there is not further independent evidence from this source. If the multiverse is to be independently credible, it will need considerable independent theoretical or empirical substantiation.

## Conclusions

In this essay, I have defended Bayesianism from a potential paradox of confirmation: the idea that trivial existential claims appear to confirm a multiverse theory according to Bayes's theorem. I noted how this can be resolved by appealing only to the most certain data—the most proximal facts—when appraising probabilities. I noted parallels between this case and the more well-known inverse gambler's fallacy and explained, likewise, how our intuitions surrounding the inverse gambler's fallacy can be accommodated by Bayesianism. All this motivates further the proximal facts approach to Bayesianism.

I ended with one exception: our own existence as embodied observers *does* appear to confirm a multiverse hypothesis. I explained how this is, in fact, consonant with our intuitions and does indeed provide the multiverse with a significant measure of support. However, this support is limited by the credibility of alternative explanations for cosmic fine-tuning and does not limit the strength of fine-tuning arguments for theism. In order to limit the strength of such arguments, substantial *independent* support for a multiverse will be required. And that is not possible simply by appealing to the diverse range of material objects and details in our universe.