

Likelihoods, Multiple Universes, and Epistemic Context

LYDIA MCGREW
Kalamazoo, Michigan

The life-permitting values of the fundamental constants in our universe puzzle physicists and philosophers alike and provide the lynchpin for the fine-tuning argument for design. But both advocates and opponents of the fine-tuning argument treat multiple universes with a selection effect as an entirely respectable hypothesis to account for the constant values. I will argue that, except where there is relevant prior information, the existence of multiple instances of a low-likelihood causal process cannot legitimately be used as an alternative to a different, higher-likelihood causal process. This conclusion is relevant to the fine-tuning debate but also has interest for those who hold, as I do, that the fine-tuning argument is flawed in any event.

I have argued elsewhere that the fine-tuning argument for a designed universe does not work because there is no well-defined probability for the values of the constants on the hypothesis of no design or “chance.”¹ This view renders the postulation of multiple undesigned universes moot, as multiplying an undefined probability many times yields an undefined probability. Furthermore, I have argued elsewhere that the action of an agent should not automatically be treated as ranging over the entirety of a group of trials, even when those trials are independently known to exist, and that a larger group of trials is only relevant to a comparative inference when both hypotheses are equally relevant to the entire group of trials.² This consideration casts doubt on the relevance of multiple universes in the fine-tuning argument, since design by an agent is one of the alternatives.

¹ Lydia McGrew, Timothy McGrew, and Eric Vestrup, “Probabilities and the Fine-Tuning Argument: A Sceptical View,” *Mind* 110 (2001): 1027–37. See also Lydia McGrew and Timothy McGrew, “On the Rational Reconstruction of the Fine-Tuning Argument: A Response to Robin Collins and Alexander R. Pruss,” *Philosophia Christi* 7 (2005): 425–443.

² Lydia McGrew, “Agency and the Metalottery Fallacy,” *Australasian Journal of Philosophy* 80 (2002): 440–64.

However, for my purposes here I intend to ignore considerations specific to universes or to design and to discuss the issue as a probabilistic problem that arises in comparing two causal hypotheses as putative explanations of an event. I shall grant for the sake of the argument here that if there are multiple trials, both hypotheses may be assumed to be equally relevant to all of them, and I shall stipulate that the likelihoods are well-defined.

Let us imagine two causal processes—say, two machines—that produce balls. Let E be the statement, “I see a black ball,” A be, “Machine A has produced a ball” and B be, “Machine B has produced a ball.” Let $P(E|A) = .001$ and $P(E|B) = 1$. Suppose that a “selection effect” is in place—that is, that a black ball is the only kind of ball I can see, so that the production of at least one black ball is a condition for the possibility of my seeing any ball. Finally, suppose that the production of a black ball is a sufficient condition for my seeing a ball.

The literature on fine-tuning contains much interesting debate about the relation of a single trial of a process having low likelihood to many trials of that same process—in terms of the ball-producing machines, the relation of A to an alternative A' , something like “Machine A has produced 10,000 balls.”³ The result is that apparently, given a selection effect, A' can reasonably be inferred over A if one finds oneself observing a black ball, at least when there are well-defined prior probabilities for the two hypotheses, and especially when these prior probabilities are equal.⁴ The result is a consequence of the selection effect condition, which means both that I cannot see any nonblack ball and therefore do not have independent access to the outcome of a given trial and also that a black ball produced anywhere among the many trials in A' brings about the result E .⁵ The arguments about this

³ Ian Hacking, “Coincidences: Mundane and Cosmological,” in *Origins and Evolution of the Universe: Evidence for Design?* ed. J. M. Robson (Montreal, 1987), 119–38; John Leslie, “Anthropic Explanations in Cosmology,” *Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1 (1986):87–95; Ian Hacking, “The Inverse Gambler’s Fallacy,” *Mind* 96 (1987): 331–40; John Leslie, “No Inverse Gambler’s Fallacy in Cosmology,” *Mind* 97 (1988): 269–72; John Leslie, *Universes* (New York: Routledge, 1989), 142–4.

⁴ It may be objected that the existence of an undesigned universe is not the same thing as some universe’s having been produced by an unintelligent causal process. I agree. But the intuition pumps used by those who argue that multiple universes are relevant *do* involve stochastic trials of chance processes—e.g., dice being rolled and my being awakened only if double six comes up. The problems with using such analogies reflect the problems with using probabilities in this area.

⁵ I have been persuaded, albeit reluctantly, that Leslie is correct and Hacking incorrect on this specific point. However, in the absence of defined priors as between A and A' , there is no reason to conclude A' even under the other conditions specified. Leslie is somewhat cavalier about prior probabilities (*Universes*, 17, 203–4). Some of the debate over multiple universes concerns whether they are correctly described in these terms, e.g., whether the outcome anywhere in the series produces the relevant result. See Roger White, “Fine Tuning and Multiple Universes,” in *God and Design*, ed. Neil A. Manson (New York: Routledge, 2003), 229–50.

issue have to some extent obscured the fact that, in the end, the point is not to compare A' and A but to compare A' and B. It is apparently taken for granted that if we could be justified by likelihoods in preferring A' to A, it is epistemically legitimate to treat A' as a serious alternative to B.

A clue that this may not be a justified assumption arises from the apples-to-oranges nature of comparing A' to B. A', after all, is the hypothesis that machine A has operated many times, whereas B is the hypothesis that machine B has produced a ball. This disparity raises the suspicion that A' has only been brought into consideration in the first place in order to try to produce a probability for E on *some* version of A that is closer to that on B—in other words, that A' is *ad hoc*.

It does, however, seem difficult to bring home an objection to the comparison of A' and B if $P(E|B)$ is very high, and especially in the limiting case where it is equal to 1. After all, if the probability of E on B is already 1, it may seem that B has already been given, as it were, all that it has coming. What more could an advocate of B ask for? In particular, it may seem that there is no point in considering the possibility of multiple trials of B—an idea to which I shall return momentarily.

But if the objection to A' is that it is *ad hoc*, then likelihoods may not be the only relevant consideration. It is no news to epistemologists and philosophers of science that one can get an arbitrarily high probability for any evidence by designing an hypothesis for that purpose. To use Elliott Sober's example, gremlins bowling in the attic—if they are the right kind of gremlins with the right kind of bowling balls—may make exactly the same sounds as squirrels running around in the attic.⁶ If I discover that there are cookies missing from the cookie jar, one explanation is that my child has taken them. But I can, if I wish, conjecture that a fairy has stolen the cookies, and this, too, will give a very high probability (at least equal to that on the hypothesis that my child is the culprit) to the fact of the missing cookies.

The analysis of *ad hocness* is a vexed issue, but we may be able usefully to pursue various heuristic approaches to detecting *ad hoc* hypotheses and, at least for the moment, leave aside the attempt to state necessary and sufficient conditions for *ad hocness*. I propose the modest claim that some *ad hoc* hypotheses include a degree of elaboration that is present only to make the conditional probability of E on the *ad hoc* hypothesis match (or to some substantial degree approach) that on some rival hypothesis. The elaboration can be repeated if further anomalous evidence arises. It may, then, be heuristically helpful for detecting some *ad hoc* hypotheses if we look for a repeatable elaboration move. In the case of the missing cookies, the first two hypotheses under consideration—child thief or fairy thief—are both fairly straightforward, although one has (under ordinary circumstances) a far

⁶ Elliott Sober, *Philosophy of Biology*, 2nd ed. (Boulder, CO: Westview, 2000), 32.

higher prior probability than the other. It may, however, occur to me that I have never seen a fairy in the kitchen. To deal with *this* evidence, the initial hypothesis can easily be elaborated: This is, I may speculate, an invisible fairy. And if that move is permissible, I can add other properties to the fairy to accommodate any new evidence that has a low probability on the invisible fairy hypothesis and a higher probability on the child-thief hypothesis. If, for example, no more cookies disappear after I put the cookie jar out of my child's reach, this result would seem to favor the child-thief hypothesis over the invisible fairy hypothesis. But, of course, the only reason the invisible fairy was invoked in the first place was because it gave a relatively high probability to the evidence already in hand. And if such considerations made it legitimate to take *that* version of the fairy theory seriously, it is hard to find any principled objection to elaborating it again: Not only is this fairy invisible, it is also vertically challenged, unable to reach higher than about four feet off the ground.

This heuristic is fairly easy to apply to A'. The obviously parallel hypothesis to B is not A' but A. A is a straightforward causal hypothesis, as is B, and the two of them are under consideration as rival causal explanations for E. A' is taken seriously precisely and solely because A gives a quite low probability to E but, granting a selection effect, A' gives a much higher probability. (The probability of at least one black ball somewhere or other in the set of 10,000 is greater than 99.99 percent.) This is why so much debate in the literature focuses on the comparison of A to A'. But A' involves postulating many rolls rather than one. The construction of A' is the initial elaboration move.

Once we are considering A', we may reflect profitably that the obviously parallel hypothesis to A' is not B but B'—"Machine B has produced 10,000 balls." A' and B' give significantly different likelihoods to *other* evidence E'—"10,000 black balls have been produced." If we were to gain evidence, perhaps indirectly, for E', this would confirm B' overwhelmingly over A', on which it would have a vanishingly small probability. But if A' was legitimately compared to B, in the absence of independent evidence favoring it, on the grounds that it gave a much higher probability to E than did A, we could elaborate A' upon obtaining E', thereby obtaining A'', "Machine A has produced *n* balls," where *n* is some far greater number of balls, sufficient to give a probability of more than 99.99 percent that 10,000 black balls will appear somewhere in the still larger group. Now A'' will be the competitor to B'. And this elaboration procedure can be repeated no matter how many black balls we find.

In the fine-tuning literature, we are generally encouraged to imagine world ensembles in which most of the universes do not have life-permitting

constants. For example, John Leslie describes a multiple universe scenario in this fashion:

[O]ne way of accounting for fine tuning of the world's properties to suit Life's needs would be [to] suppose that there exists an ensemble of vastly many 'Worlds' or 'universes' with very varied properties. Ours would be one of the rare ones in which living beings could evolve.⁷

In terms of balls, this amounts to imagining scenarios in which many balls have been produced and most of them are not black. But the claim that black balls are rare in any larger set of balls produced is not an assumption to which anyone is entitled without evidence. Of course, A' predicts a low frequency of black balls within a larger set, but B' does not. In terms of universes, this amounts to pointing out that, even if there are many universes, we cannot rule out *a priori* the possibility that many or all of them have life-permitting constants. Even independent evidence for many universes would not be *ipso facto* evidence for a theory alternative to design in the absence of evidence about the properties of the larger set. In terms of the ball machines, since A and B make different predictions about E, A' and B' make different predictions about E'. So it becomes important to separate the hypothesis that there have been many balls produced from a claim about their properties. What is sauce for the A goose is sauce for the B gander. If proponents of A are allowed to imagine a set consisting almost entirely of nonblack balls, proponents of B are allowed to imagine a set consisting entirely of black balls. To assume one or the other without independent evidence would be to beg the question at the level of the larger set.

Richard Dawkins makes exactly this sort of illicit assumption about outcome frequency when discussing the origins of life:

Begin by giving a name to the probability, however low it is, that life will originate on any randomly designated planet of some particular type. Call this number the spontaneous generation probability or SGP. . . . Suppose that our best guess of the SGP is some very very small number, say one in a billion. . . . Yet if we assume, as we are perfectly entitled to do for the sake of argument, that life has originated only once in the universe, it follows that we are *allowed* to postulate a very large amount of luck in a theory, because there are so many planets in the universe where life could have originated.⁸

⁷ Leslie, *Universes*, 6. The reference to the properties as "varied" is a bit puzzling, as it may be taken to mean that on this hypothesis the properties *must* vary, which would make universe-generating events stochastically dependent on each other. But Leslie's other analogies (throwing dice, monkeys at typewriters, etc.) make it clear that he means only to imagine that the outcomes *in fact* vary in the different trials. Stochastic independence is important in the debate with Hacking. See Hacking, "Coincidences," 134–5.

⁸ Richard Dawkins, *The Blind Watchmaker* (New York: Norton, 1987), 144.

The existence of multiple planets, unlike that of multiple trials of process A or multiple universes, is not postulated without independent evidence. But why is Dawkins entitled to assume “for the sake of the argument” that life has originated with very low frequency—only once—in the universe? In the context, he is discussing naturalistic hypotheses regarding the origins of life; the alternative under consideration is a design hypothesis. Under these circumstances, it seems that Dawkins must be assuming that life has originated only once because he is assuming that it would only originate by way of some low-probability process. Since the cause of the origin of life is the very point at issue, this is question-begging. Even when we know independently that there have been many trials, we need independent evidence as to the frequency of the outcome of interest within that set. We cannot assume a low frequency if one of the causes under consideration, but not the other, predicts a low frequency.

But if we prescind from unsupported assumptions about the composition of a larger set of balls, we are left without any way to choose between A' and B'. Would it not, then, make most sense to focus instead on the pair of hypotheses for which we do have relevant discriminating evidence—A and B? We cannot compare A' to B solely on the grounds that they both give a fairly high probability to E. For, again, A can by such elaborations keep pace with B on likelihood considerations alone no matter what evidence comes to light—an epistemic picture that is unsatisfying, to put it mildly.

The moral of the story, I submit, is that we cannot do our most important epistemic work using likelihoods alone. This is itself not entirely startling news to epistemologists, but its relevance to the many trials/many universes issue does not seem to have been recognized.

Consider how different the epistemic situation would look if we had specific relevant prior evidence. We might have strong independent evidence for A' (not merely for the claim that there *are* many balls, but that many balls have been produced by machine A), making A' not simply an elaboration of A for the sake of accommodating E. And in that case, the specter of infinite iterations would not arise. If we were to discover E', it might well be the case that we would have no comparable independent evidence supporting A''. Similarly, if we had strong independent evidence that A, A', and B were all equally likely, or that both A and A' were a great deal more probable than B, these comparative probabilities would help to answer the charge that elaborating A is an *ad hoc* move that can be repeated *ad infinitum*. For there would be no guarantee that similar independent evidence would pertain to A', A'', and B'.⁹

⁹ It is important, however, to avoid very strong off-the-cuff estimates even of prior probabilities, as these may blind us to the impact of new evidence just as surely as *ad hoc* postulates. If we merely say something like, “Oh, A is *vastly* more probable than B, so much so that I'm sure A' is also vastly more probable than B, and so much so that it can easily accommodate the apparent anomaly of E,” we may casually say the same for E' or for any evidence whatsoever.

In the absence of such data, the only point in comparing the probability of E on A and B arises from the parallels between them. The value of this comparison is lost if, arbitrarily, one hypothesis is elaborated so as to yield a likelihood comparable to that of the other. In fact, if we ignore the importance of comparing apples with apples and focus only on likelihoods—using whatever hypotheses we happen to come up with—we lose the force of comparative likelihoods themselves. There is nothing enlightening about starting out with two causal hypotheses, one of which gives a far higher probability to the evidence than the other, and then washing out that difference by “beefing up” the low-likelihood hypothesis. This maneuver shows only that we can conceive of some version of the initially low-likelihood hypothesis that yields a higher number; it demonstrates our ingenuity, but it tells us nothing about what rational credibilities we should give to the hypotheses on the basis of the evidence.

Likelihoods do the most epistemic good when they are considered in the proper epistemic context. Ultimately, we can only supply this context fully if we have priors, even fuzzy ones. If prior probabilities are missing and we can make only a likelihood comparison, it is particularly urgent for us to provide *some* rationale for comparing a pair or set of hypotheses. If we do not do so, we are merely playing with numbers which may or may not be of any epistemic consequence.