

# The Anthropic Principle Does Not Support Supernaturalism

by Michael Ikeda and Bill Jefferys

---

## Introduction

It has been claimed, most prominently by Dr. Hugh Ross on his web site

<http://www.reasons.org/resources/papers/design.html>

that the so-called "fine-tuning" of the constants of physics supports a supernatural origin of the universe. Specifically, it is claimed that many of the constants of physics must be within a very small range of their actual values, or else life could not exist in our universe. Since it is alleged that this range is very small, and since our very existence shows that our universe has values of these constants that *would* allow life to exist, it is argued that the probability that our universe arose by chance is so small that we must seek a supernatural origin of the universe.

In this article we will show that this argument is wrong. Not only is it wrong, but in fact we will show that the observation that the universe is "fine-tuned" in this sense can only count against a supernatural origin of the universe. And we shall furthermore show that with certain theologies suggested by deities that are both inscrutable and very powerful, the more "finely-tuned" the universe is, the more a supernatural origin of the universe is undermined.

[**Note added 020106:** We have learned that the philosopher of science, Elliott Sober, has made some similar points in a recent article written for the *Blackwell Guide to Philosophy of Religion*. A recent copy can be obtained [here](#): We have some slight differences with Professor Sober (in particular, we think that his condition (A3) is too strong, and that a weaker version of (A3) actually gives a stronger result), but he has an excellent discussion of the role that selection bias plays where the bias is due to self-selection by sentient observers.]

Our basic argument starts with a few very simple assumptions. We believe that anyone who accepts that the universe is "fine-tuned" for life would find it difficult not to accept these assumptions. They are:

- a) Our universe exists and contains life.
- b) Our universe is "life friendly," that is, the conditions in our universe (such as physical laws, etc.) permit or are compatible with life existing naturalistically.
- c) Life cannot exist in a universe that is governed solely by naturalistic law unless that universe is "life-friendly."

In this FAQ we will discuss only the Weak Anthropic Principle (WAP), since it is uncontroversial and generally accepted. We will not discuss the Strong Anthropic Principle (SAP), much less the Completely Ridiculous Anthropic Principle :-)

According to the WAP, which is embodied in assumption (c), the fact that life (and we as intelligent life along with it) exists in our universe, coupled with the assumption that the universe is governed by naturalistic law, implies that those laws must be "life-friendly." If they were not "life-friendly," then it is obvious that life could not exist in a universe governed solely by naturalistic law. However, it should be noted that a sufficiently powerful supernatural principle or entity (deity) could sustain life in a universe with laws that are not "life-friendly," simply by virtue of that entity's will and power.

We will show that if assumptions (a-c) are true, then the observation that our universe is "life-friendly" can *never* be evidence *against* the hypothesis that the universe is governed solely by naturalistic law. Moreover, "fine-tuning," in the sense that "life-friendly" laws are claimed to represent only a very small fraction of possible universes, can even undermine the hypothesis of a supernatural origin of the universe; and the more "finely-tuned" the universe is, the more this hypothesis can be undermined.

## Traditional responses to the "fine-tuning" argument

There are a number of traditional arguments that have been made against the "fine-tuning" argument. We will state them here, and we think that they are valid, although our main interest will be directed towards some new insights arising from a deeper understanding of probability theory.

1) In proving our main result, we do not assume or contemplate that universes other than our own exist (e.g., as in cosmologies such as those proposed by A. Vilenkin ["Quantum creation of the universe," *Phys Rev D* Vol. **30**, pp. 509-511 (1984)], André Linde ["The self-reproducing inflationary universe," *Scientific American*, November 1994, pp. 48-55], and most recently, Lee Smolin [*Life of the Cosmos*, Oxford University Press (1997)], or as in some kinds of "many worlds" quantum models). One argument against Ross has been to claim that there may be many universes with many different combinations of physical constants. If there are enough of them, a few would be able to support life solely by chance. It is hypothesized that we live in one of those few. Thus, this argument seeks to overcome the low probability of having a universe with life in it with a multiplicity of universes. A recent technical discussion of this idea by Garriga and Vilenken can be found at <http://xxx.lanl.gov/abs/gr-qc/0102010>.

2) Others have argued against the assumption that the universe must have very narrowly constrained values of certain physical constants for life to exist in it. They have argued that life could exist in universes that are very different from ours, but it is only our insular ignorance of the physics of such universes that misleads us into thinking that a universe must be much like our own to sustain life. Indeed, virtually nothing is known about the possibility of life in universes that are very different from ours. It could well be that most universes could support life, even if it is of a type that is completely unfamiliar to us. To assert that only universes very like our own could support life goes well beyond anything that we know today.

Indeed, it might well be that a fundamental "theory of everything" in physics would predict that only a very narrow range of physical constants, or even no range at all, would be possible. If this turns out to be the case, then the entire "fine-tuning" argument would be moot.

While recognizing the force and validity of these arguments, the main points we will make go in quite different directions, and show that *even if Ross is correct about "fine-tuning"* and *even if ours is the only universe that exists*, the "fine-tuning" argument fails.

## Notation and some basic probability theory

In this section, we will introduce some necessary notation and discuss some basic probability theory needed in order to understand our points

First, some notation. We introduce several predicates, (statements which can have values true or false).

Let  $L$ ="The universe exists and contains Life."  $L$  is clearly true for our universe (assumption a).

Let  $F$ ="The conditions in the universe are 'life-Friendly,' in the sense described above." Ross, in his arguments, certainly assumes that  $F$  is true. So will we (assumption b). The negation,  $\sim F$ , would be that the conditions are such that life cannot exist naturalistically, so that if life is present it must be because of some supernatural principle or entity.

Let  $N$ ="The universe is governed solely by Naturalistic law." The negation,  $\sim N$ , is that it is not governed solely by naturalistic law, that is, some non-naturalistic (supernaturalistic) principle or entity is involved.  $N$  and  $\sim N$  are not assumptions; they are hypotheses to be tested. However, we do not rule out either possibility at the outset; rather, we assume that each of them has some non-zero *a-priori* probability of being true.

Probability theory now allows us to write down some important relationships between these predicates. For example, assumption (c) can be written mathematically as  $N \& L \implies F$  (' $\implies$ ' means logical implication). In the language of probability theory, this can be expressed as

$$P(F|N \& L) = 1$$

where  $P(A|B)$  is the probability that A is true, given that B is true [see footnote 1 for a formal mathematical definition], and '&' is logical conjunction.

## Why the "fine-tuning" argument is invalid

Expressed in the language of probability theory, we understand the "fine-tuning" argument to claim that if naturalistic law applies, then the probability that a randomly-selected universe would be "life-friendly" is very small, or in mathematical terms,  $P(F|N) \ll 1$ . Notice that this condition is not a predicate like L, N and F; Rather, it is a statement about the *probability distribution*  $P(F|N)$ , considered as it applies to all possible universes. For this reason, it is not possible to express the "fine-tuning" condition in terms of one of the arguments A or B of a probability function  $P(A|B)$ . It is, rather, a statement about how large those probabilities are.

The "fine-tuning" argument then reasons that if  $P(F|N) \ll 1$ , then it follows that  $P(N|F) \ll 1$ . In ordinary English, this says that if the probability that a randomly-selected universe would be life-friendly (given naturalism) is very small, then the probability that naturalism is true, given the observed fact that the universe is "life-friendly," is also very small. This, however, is an elementary if common blunder in probability theory. One cannot simply exchange the two arguments in a probability like  $P(F|N)$  and get a valid result. A simple example will suffice to show this.

---

### Example

Let A="I am holding a Royal Flush."

Let B="I will win the poker hand."

It is evident that  $P(A|B)$  is nearly 0. Almost all poker hands are won with hands other than a Royal Flush. On the other hand, it is equally clear that  $P(B|A)$  is nearly 1. If you have a Royal Flush, you are virtually certain to win the poker hand.

---

There is a second reason why this "fine-tuning" argument is wrong. It is that for an inference to be valid, it is necessary to take into account *all* known information that may be relevant to the conclusion. In the present case, we happen to *know* that life exists in our universe (i.e., that L is true). Therefore, it is invalid to make inferences about N if we fail to take into account the fact that L, as well as F, are already known to be true. It follows that any inferences about N *must* be conditioned upon *both* F *and* L. An example of this is seen in the next section.

The most important consequence of the previous paragraph is very simple: In inferring the probability that N is true, it is entirely irrelevant whether  $P(F|N)$  is large or small. It is entirely irrelevant whether the universe is "fine-tuned" or not. Only probabilities conditioned upon L are relevant to our inquiry.

Richard Harter <cri@tiac.net> has suggested a somewhat different interpretation of the "fine-tuning" argument in E-mail (reproduced here with permission). He writes:

This takes care of the WAP; if one argues solely from the WAP the FAQ argument is correct. However the "fine tuning" argument is not (despite what its proponents say) a WAP argument; it is an inverse Bayesian argument. The argument runs thusly:

$$P(F|\sim N) \gg P(F|N)$$

ergo

$$P(\sim N|F) \gg P(N|F)$$

Considered as a formal inference this is a fallacy. None-the-less it is a normal rule of induction which is (usually) sound. The reason is that for the "conclusion" not to hold we need

$$P(N) \gg P(\sim N)$$

[This is not the full condition but it is close enough for government work.]

There are two fallacies in this form of the argument. The first is the failure to condition on L, mentioned above. This in itself would render the argument invalid. The second is that the first line of the argument,  $P(F|\sim N) \gg P(F|N)$ , is merely an unsupported assertion. No one knows what the probability of a supernatural entity creating a universe that is F is! For example, a dilettante deity might never get around to creating any universes at all, much less ones capable of supporting life.

[**Note added 010612:** Since this was written, we have proved that if You, knowing as a sentient observer that L is true, adopt an *a priori* position that is neutral between N and  $\sim N$ , i.e., that  $P(\sim N|L)$  is of the same order of magnitude as  $P(N|L)$ , then when You learn that F is true *and* that  $P(F|N) \ll 1$ , You will conclude that  $P(F \& L \& \sim N) \ll 1$ . See Appendix 1 (Reply to Kwon) at the end of this essay for the proof. This observation is problematic for Harter's argument. For under these assumptions we have

$$P(F \& L \& \sim N) = P(L|F \& \sim N)P(F|\sim N)P(\sim N) \ll 1.$$

Thus under these assumptions it follows that at least one of  $P(L|F \& \sim N)$ ,  $P(F|\sim N)$  or  $P(\sim N)$  is quite small. A small  $P(L|F \& \sim N)$  says that it is almost certain that the supernatural deity, having created a "life-friendly" universe, would make it sterile (lifeless). A small  $P(F|\sim N)$  says that it is highly *unlikely* that this deity would even create a universe that is "life-friendly". Both of these undermine the usual concepts attributed to the deity by "intelligent design" theorists, although either would be consistent with a deity that was incompetent, a dilettante, or a "trickster". A small  $P(F|\sim N)$  is also consistent with a deity who makes many universes, most of them being  $\sim F$ , with many of these  $\sim F$  universes perhaps containing life (that is,  $\sim F \& L$  universes, as we discuss below). A small  $P(\sim N)$  says that it is nearly certain that naturalism is true *a priori* and unconditioned on L, so that Harter's "escape" condition  $P(N) \gg P(\sim N)$  in fact holds.

Please remember that if You are a sentient observer, You must already know that L is true, even before You learn anything about F or  $P(F|N)$ . Thus it is legitimate, appropriate, and indeed *required*, for You to elicit Your prior on N versus  $\sim N$  conditioned on L and use that as Your starting point. If You then retrodict that  $P(\sim N) \ll 1$  as a consequence, all You are doing is eliciting the prior that You would have had in the absence of Your knowledge that You existed as a sentient observer. This is the only legitimate way to infer Your value of  $P(\sim N)$  unconditioned on L.]

## Our main theorem

Having understood the previous discussion, and with our notation in hand, it is now easy to prove that the WAP does not support supernaturalism (which we take to be the negation  $\sim N$  of N). Recall that the WAP can be written as  $P(F|N \& L) = 1$ . Then, by Bayes' theorem [see footnote 2] we have

$$\begin{aligned} P(N|F \& L) &= P(F|N \& L)P(N|L)/P(F|L) \\ &= P(N|L)/P(F|L) \\ &\gg P(N|L) \end{aligned}$$

where '>=' means "greater than or equal to." The second line follows because  $P(F|N\&L)=1$ , and the inequality of the third line follows because  $P(F|L)$  is a positive quantity less than or equal to 1. (The above demonstration is inspired by a recent article on talk.origins by Michael Ikeda <mmikeda@erols.com>; we have simplified the proof in his article. The message ID for the cited article is <5j6dq8\$bvj@winter.erols.com> for those who wish to search for it on dejanews.)

The inequality  $P(N|F\&L) \geq P(N|L)$  shows that the WAP supports (or at least does not undermine) the hypothesis that the universe is governed by naturalistic law. This result is, as we have emphasized, independent of how large or small  $P(F|N)$  is. The observation  $F$  cannot decrease the probability that  $N$  is true (given the known background information that life exists in our universe), and may well increase it.

Corollary: Since  $P(\sim N|F\&L)=1-P(N|F\&L)$  and similarly for  $P(\sim N|L)$ , it follows that  $P(\sim N|F\&L) \leq P(\sim N|L)$ . In other words, the observation  $F$  does not support supernaturalism ( $\sim N$ ), and may well undermine it.

### Another way to look at it

The thrust of practically all "Intelligent Design" and Creationist arguments (excepting the anthropic argument and perhaps a few others) is to show  $\sim F$ , since it is evident, we think, that if  $\sim F$  then we cannot have both life and a naturalistic universe. We evidently do have life, so the success of one of these arguments would clearly establish  $\sim N$ . In other words, given our prior opinion  $P(N\&L)$ , where  $0 < P(N\&L) < 1$  but otherwise unrestricted (thus we neither rule in nor rule out  $N$  initially), arguments like Behe's attempt to support  $\sim F$  so as to undermine  $N$ :

$$P(N|\sim F\&L) < P(N|L).$$

But the "anthropic" argument is that observing  $F$  also undermines  $N$ :

$$P(N|F\&L) < P(N|L).$$

We assert that the intelligent design folks want these inequalities to be strict (otherwise there would be no point in their making the argument!)

From these two inequalities we readily derive a contradiction, as follows. From the definition of conditional probability [see footnote 1], the two inequalities above yield

$$P(N\&\sim F\&L) < P(N|L)P(\sim F\&L), \quad P(N\&F\&L) < P(N|L)P(F\&L)$$

Adding,

$$\begin{aligned} P(N\&L) &= P(N\&\sim F\&L) + P(N\&F\&L) \\ &< P(N|L)(P(\sim F\&L) + P(F\&L)) \\ &= P(N|L)P(L) = P(N\&L), \end{aligned}$$

a contradiction since the inequality is strict.

If we remove the restriction that the inequalities be strict, then the only case where both inequalities can be true is if

$$P(N|\sim F\&L) = P(N|L) \quad \text{and} \quad P(N|F\&L) = P(N|L).$$

In other words, the only case where both can be true is if the information that the universe is "life-friendly" has *no* effect on the probability that it is naturalistic (given the existence of life); and this can only be the case if neither inequality is strict.

In essence, we see that the intelligent design folks who make the anthropic argument are really trying to have it both ways: They want observation of F to undermine N, and they also want observation of  $\sim F$  to undermine N. That is, they want any observation whatsoever to undermine N! But the error is that the anthropic argument does *not* undermine N, it supports N. They can have one of the prongs of their argument, but they can't have both.

[**Note added 010612:** Some people have objected to us that Behe is not making the argument  $\sim F$ , but is only making a statement that it is *highly unlikely* that certain of his "IC" structures could arise naturalistically. Our reading of Behe that he is making an argument that it is *impossible* for this to happen (a form of  $\sim F$  as we understand it), but even if we are wrong and he is not making this argument, the point of our comments in this section is that making the argument that the universe is F or is "fine-tuned" ( $P(F|N) \ll 1$ ) does not support supernaturalism; the argument that should be made is that the universe is  $\sim F$ , since this manifestly supports supernaturalism by refuting naturalism. See Appendix 1 (Reply to Kwon) at the end of this essay.]

## Implications of "fine-tuning" versus mere "life-friendliness"

Ross' argument discusses the case where the conditions in our universe are not only "life-friendly," but they are also "fine-tuned," in the sense that only a very small fraction of possible universes can be "life-friendly." We have shown that regardless how "finely-tuned" the the laws of physics are, the observation that the universe is capable of sustaining life cannot undermine N.

As we have pointed out above, others have responded to the claim of "fine-tuning" in several ways. One way has been to point out that this claim is not corroborated by any theoretical understanding about what forms of life might arise in universes with different physical conditions than our own, or even any theoretical understanding about what kinds of universes are possible at all; it is basically a claim founded upon our own ignorance of physics. To those that make this point, the argument is about whether  $P(F|N)$  is really small (as Ross claims), or is in fact large. The point (against Ross) is essentially that Ross' crucial assumption is completely without support.

A second response is to point out that several theoretical lines of evidence indicate that many other, and perhaps even an infinite number of other universes, with varying sets of physical constants and conditions, might well exist, so that even if the probability that a given universe would have constants close to those of our own universe is small, the sheer number of such universes would virtually guarantee that *some* of them would possess constants that would allow life to arise.

Nevertheless, it is necessary to consider the implications of Ross' assertion that the universe is "fine-tuned." Suppose it is true that amongst all naturalistic universes, only a very small proportion could support life. What would this imply?

We have shown that the WAP tends to support N, and cannot undermine it. This observation is *independent* of whether  $P(F|N)$  is small or large, since (as we have seen) the only probabilities that are significant for inference about N are those that are conditioned upon all relevant data at our disposal, including the fact that L is true. Therefore, regardless of the size of  $P(F|N)$ , valid reasoning shows that observing that F is true cannot decrease the probability that N is true, and may increase it.

We believe that the real import of observing that  $P(F|N)$  is small (if indeed that is true) would be to strengthen Vilenkin/Linde/Smolin-type hypotheses that multiple universes with varying physical constants may exist. If indeed the universe is governed by naturalistic laws, and if indeed the probability that a universe governed by naturalistic laws can support life is small, then this supports a Vilenkin/Linde/Smolin model of multiple universes over a model that includes only a single universe with a single set of physical constants.

To see this, let S="there is only a Single universe," and M="there are Multiple universes." Let E = "there Exists a universe with life." Clearly,  $P(E|N) < P(F|N)$ , since it is possible that a universe that is "life-friendly" could still be barren. But, since L is true, E is also true, so observing L implies that we have also observed E.

Then, assuming that  $P(F|N) < 1$  is the probability that a single universe is "life-friendly," that this probability is the same for each "random" multiple universe as it would be for a single universe, and that the probability that a given universe exists is independent of the existence of other universes, it follows that

$P(E|S\&N) = p = P(E|N) < P(F|N) < 1$  (and for Ross,  $P(F|N) \ll 1$ );

$P(E|M\&N) = 1 - (1-p)^m$ , where  $m$  is the number of universes if  $M$  is true; This is less than 1 but approaches 1 (for fixed  $p$ ) as  $m$  gets larger and larger. Since all the Multiple-universe proposals we have seen suggest that  $m$  is in fact infinite, it follows that  $P(E|M\&N)=1$ . (If one postulates that  $m$  is finite, then the calculation depends explicitly on  $p$  and  $m$ ; this is left as an exercise for the reader.)

Since

$$P(S|E\&N) = P(E|S\&N)P(S|N)/P(E|N) \text{ and}$$

$$P(M|E\&N) = P(E|M\&N)P(M|N)/P(E|N),$$

with these assumptions it follows by division that

$$\frac{P(M|E\&N)}{P(S|E\&N)} = \frac{1}{p} \times \frac{P(M|N)}{P(S|N)},$$

which shows that observing  $E$  (or  $L$ ) increases the evidence for  $M$  against  $S$  in a naturalistic universe by a factor of at least  $1/p$ . The smaller  $P(F|N)=p$  (that is, the more "finely-tuned" the universe is), the more likely it is that some form of multiple-universe hypothesis is true.

## Theological considerations

The next section is rather more speculative, depending as it does upon theological notions that are hard to pin down, and therefore should be taken with large grains of salt. But it is worth considering what effect various theological hypotheses would have on this argument. It is interesting to ask the question, "given that observing that  $F$  is true cannot undermine  $N$  and may support it, by how much can  $N$  be strengthened (and  $\sim N$  be undermined) when we observe that  $F$  is true?"

It is evident from the discussion of the main theorem that the key is the denominator  $P(F|L)$ . The smaller that denominator, the greater the support for  $N$ . Explicitly we have

$$P(F|L) = P(F|N\&L)P(N|L) + P(F|\sim N\&L)P(\sim N|L)$$

But since  $P(F|N\&L)=1$  we can simplify this to

$$P(F|L) = P(N|L) + P(F|\sim N\&L)P(\sim N|L).$$

Plugging this into the expression  $P(N|F\&L) = P(N|L)/P(F|L)$  we obtain

$$\begin{aligned} P(N|F\&L) &= P(N|L) / [P(N|L) + P(F|\sim N\&L)P(\sim N|L)] \\ &= 1 / [1 + P(F|\sim N\&L)P(\sim N|L) / P(N|L)] \\ &= 1 / [1 + C P(F|\sim N\&L)], \end{aligned}$$

where  $C = P(\sim N|L) / P(N|L)$  is the prior odds in favor of  $\sim N$  against  $N$ . In other words,  $C$  is the odds that we would offer in favor of  $\sim N$  over  $N$  before noting that the universe is "fine-tuned" for life.

A major controversy in statistics has been over the choice of prior probabilities (or in this case prior odds). However, for our purposes this is not a significant consideration, as long as we don't choose  $C$  in such a way as to completely rule out either possibility ( $N$  or  $\sim N$ ), i.e., as long as we haven't made up our minds in advance. This means that any positive, finite value of  $C$  is acceptable.

One readily sees from this formula that for acceptable C

(1) as  $P(F|\sim N\&L)\rightarrow 0$ ,  $P(N|F\&L)\rightarrow 1$ ;

(2) as  $P(F|\sim N\&L)\rightarrow 1$ ,  $P(N|F\&L)\rightarrow 1/[1+P(\sim N|L)/P(N|L)]=P(N|L)$ ,

where ' $\rightarrow$ ' means "approaches as a limit" and the last result follows from the fact that  $P(N|L)+P(\sim N|L)=1$ .

So,  $P(N|F\&L)$  is a monotonically decreasing function of  $P(F|\sim N\&L)$  bounded from below by  $P(N|L)$ . This confirms the observation made earlier, that noting that F is true can never decrease the evidential support for N. Furthermore, the only case where the evidential support is unchanged is when  $P(F|\sim N\&L)$  is identically 1. This is interesting, because it tells us that the only case where observing the truth of F does *not* increase the support for N is precisely the case when the likelihood function  $P(F|x\&L)$ , evaluated at F, and with x ranging over N and  $\sim N$ , cannot distinguish between N and  $\sim N$ . That is, the only way to prevent the observation F from increasing the support for N is to assert that  $\sim N\&L$  also *requires* F to be true. Under these circumstances we cannot distinguish between N and  $\sim N$  on the basis of the data F. In a deep sense, the two hypotheses represent, and in fact, *are* the same hypothesis. Put another way, to assume that  $P(F|\sim N\&L)=1$  is to concede that life in the world actually arose by the operation of an agent that is observationally indistinguishable from naturalistic law, insofar as the observation F is concerned. In essence, any such agent is just an extreme version of the "God-of-the-gaps," whose existence has been made superfluous as far as the existence of life is concerned. Such an assumption would completely undermine the proposition that it is *necessary* to go outside of naturalistic law in order to explain the world as it is, although it doesn't undermine any argument for supernaturalism that doesn't rely on the universe being "life-friendly".

So, if supernaturalism is to be distinguished from naturalism on the basis of the fact that the universe is F, it must be the case that  $P(F|\sim N\&L)<1$ . Otherwise, we are condemned to an unsatisfying kind of "God-of-the-gaps" theology. But what sort of theologies can we consider, and how would they affect this crucial probability?

To make these ideas more definite, we consider first a specific interpretation that is intended to imitate, albeit crudely, how the assumption of a relatively powerful and inscrutable deity (such as a generic Judeo-Christian-Islamic deity might be) could affect the calculation of the likelihood function  $P(F|\sim N\&L)$ .

We suggest that any reasonable version of supernaturalism with such a deity would result in a value of  $P(F|\sim N\&L)$  that is, in fact, very small (assuming that only a small set of possible universes are F). The reason is that a sufficiently powerful deity could arrange things so that a universe with laws that are not "life-friendly" can sustain life. Since we do not know the purposes of such a deity, we must assign a significant amount of the likelihood function to that possibility. Furthermore, if such a deity creates universes and if the "fine-tuning" claims are correct, then *most* life-containing universes will be of this type (i.e., containing life despite not being "life-friendly"). Thus, all other things being equal, and if this is the sort of deity we are dealing with, we would *expect* to live in a universe that is  $\sim F$ .

To assert that such a deity *could* only create universes containing life if the laws are life-friendly is to restrict the power of such a deity. And to assert that such a deity *would* only create universes with life if the laws are life-friendly is to assert knowledge of that deity's purposes that many religions seem reluctant to claim. Indeed, any such assertion would tend to undermine the claim, made by many religions, that their deity can and does perform miracles that are contrary to naturalistic law, and recognizably so.

Our conclusion, therefore, is that not only does the observation F support N, but it supports it overwhelmingly against its negation  $\sim N$ , if  $\sim N$  means creation by a sufficiently powerful and inscrutable deity. This latter conclusion is, by the way, a consequence of the Bayesian Ockham's Razor [Jefferys, W.H. and Berger, J.O., "Ockham's Razor and Bayesian Analysis," *American Scientist* **80**, 64-72 (1992)]. The point is that N predicts outcomes much more sharply and narrowly than does  $\sim N$ ; it is, in Popperian language, more easily falsifiable than is  $\sim N$ . (We do not wish to get into a discussion of the Demarcation Problem here since that is out of the scope of this FAQ, though we do not regard it as a difficulty for our argument. For our purposes, we are simply making a statement about the consequences of the likelihood function having significant support on only a relatively small subset of possible outcomes.) Under these circumstances, the Bayesian Ockham's Razor shows that observing an outcome allowed by both N and  $\sim N$  is likely to favor N over  $\sim N$ . We refer the reader to the cited paper for a more detailed discussion of this point.



Aside from sharply limiting the likely actions of the deity (either by making it less powerful or asserting more human knowledge of the deity's intentions), we can think of only one way to avoid this conclusion. One might assert that any universe with life would appear to be "life-friendly" from the vantage point of the creatures living within it, regardless of the physical constants that such a universe were equipped with. In such a case, observing F cannot change our opinion about the nature of the universe. This is certainly a possible way out for the supernaturalist, but this solution is not available to Ross because it contradicts his assertions that the values of certain physical constants *do* allow us to distinguish between universes that are "life-friendly" and those that are not. And, such an assumption does not come without cost; whether others would find it satisfactory is problematic. For example, what about miracles? If every universe with life looks "life-friendly" from the inside, might this not lead one to wonder if everything that happens therein would also look to its inhabitants like the result of the simple operation of naturalistic law? And then there is Ockham's Razor: What would be the point of postulating a supernatural entity if the predictions we get are indistinguishable from those of naturalistic law?

## But which deity?

In the previous section, we have discussed just one of many sorts of deities that might exist. This one happens to be very powerful and rather inscrutable (and is intended to be a model of a generic Judeo-Christian-Islamic sort of deity, though believers are welcome to disagree and propose--and justify--their own interpretations of their favorite deity). However, there are many other sorts of deities that might be postulated as being responsible for the existence of the universe. There are somewhat more limited deities, such as Zeus/Jupiter, there are deities that share their existence with antagonistic deities such as the Zoroastrian Ahura-Mazda/Ahriman pair of deities, there are various Native American deities such as the trickster deity Coyote, there are Australian, Chinese, African, Japanese and East Indian deities, and even many other possible deities that no one on Earth has ever thought of. There could be deities of lifeforms indigenous to planets around the star Arcturus that we should consider, for example.

Now when considering a multiplicity of deities, say  $D_1, D_2, \dots, D_i, \dots$ , we would have to specify a value of the likelihood function for each individual deity, specifying what the implications would be if *that* deity were the actual deity that created the universe. In particular, with the "fine-tuning" argument in mind, we would have to specify  $P(F|D_i \& L)$  for every  $i$  (probably an infinite set of deities). Assuming that we have a mutually exclusive and exhaustive list of deities, we see the hypothesis  $\sim N$  revealed to be *composite*, that is, it is a combination or union of the individual hypotheses  $D_i$  ( $i=1, 2, \dots$ ). Our character set doesn't have the usual "wedge" character for "or" (logical disjunction), so we will use 'v' to represent this operation. We then have

$$\sim N = D_1 \vee D_2 \vee \dots \vee D_i \vee \dots$$

Now, the total prior probability of  $\sim N$ ,  $P(\sim N|L)$ , has to be divvied up amongst all of the individual subhypotheses  $D_i$ :

$$P(\sim N|L) = P(D_1|L) + P(D_2|L) + \dots + P(D_i|L) + \dots$$

where  $0 < P(D_i) < P(\sim N|L) < 1$  (assuming that we only consider deities that might exist, and that there are at least two of them). In general, each of the individual prior probabilities  $P(D_i|L)$  would be very small, since there are so many possible deities. Only if some deities are *a priori* much more likely than others would any individual deity have an appreciable amount of prior probability.

This means that in general,  $P(D_i|L) \ll 1$  for all  $i$ .

Now when we originally considered just  $N$  and  $\sim N$ , we calculated the posterior probability of  $N$  given  $L \& F$  from the prior probabilities of  $N$  and  $\sim N$  given  $L$ , and the likelihood functions. Here it would be simpler to look at prior and posterior odds. These are derived straightforwardly from probabilities by the relation

$$\text{Odds} = \text{Probability} / (1 - \text{Probability}).$$

This yields a relationship between the prior and posterior odds of  $N$  against  $\sim N$  [using  $P(N|F \& L) + P(\sim N|F \& L) = 1$ ]:

$$\begin{aligned} \text{Posterior Odds} &= \frac{P(N|F\&L)}{P(\sim N|F\&L)} = \frac{P(F|N\&L)}{P(F|\sim N\&L)} \times \frac{P(N|L)}{P(\sim N|L)} \\ &= (\text{Bayes Factor}) \times (\text{Prior Odds}) \end{aligned}$$

The Bayes Factor and Prior Odds are given straightforwardly by the two ratios in this formula.

Since  $P(F|N\&L)=1$  and  $P(F|\sim N\&L)\leq 1$ , it follows that the posterior odds are greater than or equal to the prior odds (this is a restatement of our first theorem, in terms of odds). This means that observing that F is true cannot decrease our confidence that N is true.

But by using odds instead of probabilities, we can now consider the individual sub-hypotheses that make up  $\sim N$ . For example, we can calculate prior and posterior odds of N against any individual  $D_i$ . We find that

$$\text{Posterior Odds} = \frac{P(N|F\&L)}{P(D_i|F\&L)} = \frac{P(F|N\&L)}{P(F|D_i\&L)} \times \frac{P(N|L)}{P(D_i|L)}$$

This follows because (by footnote 2)

$$P(N|F\&L) = \frac{P(F|N\&L)P(N|L)}{P(F|L)},$$

$$P(D_i|F\&L) = \frac{P(F|D_i\&L)P(D_i|L)}{P(F|L)},$$

and the  $P(F|L)$ 's cancel out when you take the ratio.

Now, even if  $P(F|D_i\&L)=1$ , which is the maximum possible, the posterior odds against  $D_i$  may still be quite large. The reason for this is that the prior probability of  $\sim N$  has to be shared out amongst a large number of hypotheses  $D_j$ , each one greedily demanding its own share of the limited amount of prior probability available. On the other hand, the hypothesis N has no others to share with. In contrast to  $\sim N$ , which is a compound hypothesis, N is a simple hypothesis. As a consequence, and again assuming that no particular deity is *a priori* much more likely than any other (it would be incumbent upon the proposer of such a deity to explain *why* his favorite deity is so much more likely than the others), it follows that the hypothesis of naturalism will end up being much more probable than the hypothesis of *any particular* deity  $D_i$ .

This phenomenon is a second manifestation of the Bayesian Ockham's Razor discussed in the Jefferys/Berger article (cited above).

In theory it is now straightforward to calculate the posterior odds of N against  $\sim N$  if we don't particularly care *which* deity is the right one. Since the  $D_i$  form a mutually exclusive and exhaustive set of hypotheses whose union is  $\sim N$ , ordinary probability theory gives us

$$\begin{aligned} P(\sim N|F\&L) &= P(D_1|F\&L) + P(D_2|F\&L) + \dots \\ &= [P(F|D_1\&L)P(D_1|L) + P(F|D_2\&L)P(D_2|L) + \dots]/P(F|L) \end{aligned}$$

Assuming we know these numbers, we can now calculate the posterior odds of N against  $\sim N$  by dividing the above expression into the one we found previously for  $P(N|F\&L)$ . Of course, in practice this may be difficult! However, as can be seen from this formula, the deities  $D_i$  that contribute most to the denominator (that is, to the supernaturalistic hypothesis) will be the ones that have the largest values of the likelihood function  $P(F|D_i\&L)$  or the largest prior probability  $P(D_i|L)$  or both. In the first case, it will be because the particular deity is closer to predicting what naturalism predicts (as regards F), and is therefore closer to being a "God-of-the-gaps" deity; in the second, it will be because we already favored that particular deity over others *a priori*.

## Final comments

Some make the mistake of thinking that "fine-tuning" and the anthropic principle support supernaturalism. This mistake has two sources.

The first and most important of these arises from confusing entirely different conditional probabilities. If one observes that  $P(F|N)$  is small (since most hypothetical naturalistic universes are not "fine-tuned" for life), one might be tempted to turn the probability around and decide, *incorrectly*, that  $P(N|F)$  is also small. But as we have seen, this is an elementary blunder in probability theory. We find ourselves in a universe that is "fine-tuned" for life, which would be unlikely to come about by chance (because  $P(F|N)$  is small), *therefore* (we conclude *incorrectly*),  $P(N|F)$  must also be small. This common mistake is due to confusing two entirely different *conditional* probabilities. *Most* actual outcomes are, in fact, highly improbable, but it does not follow that the hypotheses that they are conditioned upon are themselves highly improbable. It is therefore fallacious to reason that if we have observed an improbable outcome, it is necessarily the case that a hypothesis that generates that outcome is itself improbable. One *must* compare the probabilities of obtaining the observed outcome under *all* hypotheses. In general, most, if not all of these probabilities will be very small, but some hypotheses will turn out to be much more favored by the actual outcome we have observed than others.

The second source of confusion is that one *must* do the calculations taking into account *all* the information at hand. In the present case, that *includes* the fact that life is known to exist in our universe. The possible existence of hypothetical naturalistic universes where life does not exist is entirely irrelevant to the question at hand, which *must* be based on the data we *actually* have.

In our view, similar fallacious reasoning may well underlie many other arguments that have been raised against naturalism, not excluding design and "God-of-the-Gaps" arguments such as Michael Behe's "Irreducible Complexity" argument (in his book, *Darwin's Black Box*), and William Dembski's "Complex Specified Information," as described in his dissertation (University of Illinois at Chicago). We conclude that whatever their rhetorical appeal, such arguments need to be examined much more carefully than has happened so far to see if they have any validity. But that discussion is outside the scope of this article.

Bottom line: The anthropic argument should be dropped. It is wrong. "Intelligent design" folks should stick to trying to undermine N by showing  $\sim F$ . That's their only hope (though we believe it to be a forlorn one).

---

Michael Ikeda	Bill Jefferys
Statistical Research Division	Department of Astronomy
Bureau of the Census	University of Texas
Washington DC 20233	Austin TX 78712

Department of Statistics  
University of Vermont  
Burlington VT

To email comments on this document, [click here](#)

Michael Ikeda's work on this article was done on his own time and not as part of his official duties. The authors' affiliations are for identification only. The opinions expressed herein are those of the authors, and do not necessarily represent the opinions of the authors' employers.

Copyright (C) 1997-2006 by Michael Ikeda and Bill Jefferys. Portions of this FAQ are Copyright (C) 1997 by Richard Harter. All Rights Reserved.

---

## Footnotes

[1] By definition,  $P(A|B)=P(A\&B)/P(B)$ ; it follows that also  $P(A|B\&C)=P(A\&B|C)/P(B|C)$ .

[2] We use Bayes' theorem in the form

$$P(A|B\&K)=P(B|A\&K)P(A|K)/P(B|K)$$

which follows straightforwardly from the identity

$$P(A|B\&K)P(B|K)=P(A\&B|K)=P(B|A\&K)P(A|K)$$

(a consequence of footnote 1) assuming that  $P(B|K)>0$ .

---

## APPENDIX 1: Reply to Kwon (April 30, 2001)

David Kwon has posted a [web page](#) in which he claims to have refuted the arguments in our article. However, he has made a simple error, which we detail below, along with comments on some of his other assertions.

[**Note added 040109**: Kwon's original article has disappeared from the web. The above link is to the last version of his article archived by the [Internet Wayback Machine](#) via [Makeashorterlink.com](#)]

Kwon's Equation (3) reads as follows:

$$P(N|F\&L) = P(N\&F\&L) / \{P(\sim N\&F\&L) + P(N\&F\&L)\}$$

This is an elementary result of probability theory and we agree with it. Kwon then goes on and assumes what he calls the "fine-tuning" condition  $P(F|N)\ll 1$  from which he correctly derives Equation (8), the important part of which reads

$$P(N\&F\&L) \ll 1$$

From these two results (3 and 8) Kwon derives

$$P(N|F\&L)\ll 1 \text{ unless } P(\sim N\&F\&L)\ll 1$$

Unfortunately, nothing in Kwon's "proof" shows that  $P(\sim N\&F\&L)$  is not  $\ll 1$ , so he cannot assert unconditionally that  $P(N|F\&L)\ll 1$  as a consequence of his assumptions. He asserts

"The only way not to come to this conclusion [that  $P(N|F\&L)\ll 1$ ] is to start with an *a priori* assumption of  $P(\sim N\&F\&L)\ll 1$ . In other words, the only way to hold on to naturalism is by assuming that theism is virtually impossible to begin with."

This, however, is incorrect, and here the "proof" falls apart. Kwon apparently recognizes that according to his Equation (3), the value of  $P(N|F\&L)$  is not governed by the actual size of  $P(N\&F\&L)$ , but instead by the relative sizes of  $P(N\&F\&L)$  and  $P(\sim N\&F\&L)$ . In particular, if  $P(N\&F\&L)\ll P(\sim N\&F\&L)$  then  $P(N|F\&L)$  will be close to zero; if  $P(N\&F\&L)$  is approximately equal to  $P(\sim N\&F\&L)$ , then  $P(N|F\&L)$  will be of order one-half; and if  $P(N\&F\&L)\gg P(\sim N\&F\&L)$ , then  $P(N|F\&L)$  will be nearly unity. Therefore, we need to look at the ratio  $R = P(N\&F\&L)/P(\sim N\&F\&L)$  to see what factors govern its size and what assumptions this entails.

We obtain:

$$R = P(N\&F\&L) / P(\sim N\&F\&L)$$

$$= \{P(F|N\&L) P(N\&L)\} / \{P(F|\sim N\&L) P(\sim N\&L)\} \quad (A)$$

$$= P(N\&L) / \{P(F|\sim N\&L) P(\sim N\&L)\} \quad (B)$$

$$\geq P(N\&L) / P(\sim N\&L) \quad (C)$$

$$= \{P(N|L) P(L)\} / \{P(\sim N|L) P(L)\} \quad (D)$$

$$= P(N|L) / P(\sim N|L) \quad (E)$$

Here, (A) and (D) follow from the definition of conditional probability, (B) by the WAP--which Kwon says he accepts--and which asserts that  $P(F|N\&L)=1$ , (C) because the probability  $P(F|\sim N\&L)$  in the denominator is  $\leq 1$ , and (E) by cancellation of  $P(L)$  in numerator and denominator.

Thus we see that in fact the ratio  $R$  cannot be small unless  $P(N|L)/P(\sim N|L)$  is also small. Therefore we cannot conclude that  $P(N|F\&L)\ll 1$  unless  $P(N|L)/P(\sim N|L)\ll 1$ --regardless of the size of  $P(N\&F\&L)$ . But what is  $P(N|L)/P(\sim N|L)$ ? Why, it is just the prior odds ratio that You assign to describe Your relative belief in  $N$  and  $\sim N$  before You learn that  $F$  is true. Thus, although Kwon is correct in noting that the only way to keep  $P(N|F\&L)$  from being very small is to have  $P(\sim N\&F\&L)\ll 1$ , this does not represent a prior commitment to naturalism as he asserts. Indeed, a prior commitment to naturalism would be to assume that  $P(N|L)/P(\sim N|L)\gg 1$ , and as (E) shows, if we assume  $P(N|L)/P(\sim N|L)$  of order unity, which reflects a neutral prior position between the  $N$  and  $\sim N$ , and not a prior commitment to naturalism, we will end up being at least neutral between  $N$  and  $\sim N$  after observing that  $F$  is true, regardless of the size of  $P(N\&F\&L)$  and  $P(F|N)$ .

Indeed, it requires a prior commitment to *supernaturalism* to get  $P(N|F\&L)\ll 1$ , because You would have to presume *a priori* that  $P(N|L)\ll P(\sim N|L)$ . Kwon has it exactly backwards.

So the absolute size of  $P(N\&F\&L)$  and  $P(F|N)$  do not tell us anything about  $P(N|F\&L)$ ; this is a confusion between conditional and unconditional probability. The only thing that counts is the ratio  $R$ . Kwon's calculation in his steps (4-8) is simply irrelevant to the final result. Indeed, we have the following theorem:

Theorem: If  $p(F|N)\ll 1$  and You are exactly neutral between  $N$  and  $\sim N$  before learning  $F$ , then  $P(\sim N\&F\&L)\ll 1$ .

Proof: Under the assumptions we have  $P(F\&N\&L)=P(N|L)P(L)\ll 1$ ; but if we are exactly neutral between  $N$  and  $\sim N$  before learning  $F$  we have  $P(N|L)=0.5=O(1)$  so the unconditional probability  $P(L)\ll 1$ . But by standard probability theory  $P(\sim N\&F\&L)\leq P(L)\ll 1$ . QED.

Thus, far from reflecting a prior commitment to naturalism as Kwon claims, the result  $P(\sim N\&F\&L)\ll 1$  is a consequence of the fine tuning condition together with the adoption of an at least neutral prior position on  $N$  versus  $\sim N$ . It is due to the fact that  $P(N\&L\&F)$  and  $P(\sim N\&L\&F)$  both have  $P(L)\ll 1$  as a factor when they are expanded using the definition of conditional probability.

Furthermore, it is even possible for  $P(\sim N|F\&L)$  to be very small (and therefore  $P(N|F\&L)$  close to unity), without making a prior commitment to naturalism. For example, suppose we adopt the neutral position  $P(N|L)=P(\sim N|L)=0.5$ ; then from (B) we find that  $R = 1/P(F|\sim N\&L)$ , and if  $P(F|\sim N\&L)\ll 1$  then  $R\gg 1$  and  $P(F|N\&L)$  is close to unity. But what does  $P(F|\sim N\&L)\ll 1$  mean? Is this a "prior commitment to naturalism?" No, a prior commitment to naturalism would involve some conditional probability on  $N$ , not some conditional probability on  $F$ . The condition  $P(F|\sim N\&L)\ll 1$  actually means that it is likely that an inhabitant of a supernaturalistically created universe would find that it is  $\sim F$ : a universe where life exists despite the fact that it could not exist naturalistically, for example as a consequence of the suspension of natural law by the supernatural creator. We discussed this extensively in our article. Indeed, without psychoanalyzing the Deity and analysing its powers and intentions, it is a priori quite likely that the Deity might create universes that are  $\sim F\&L$ , for such universes are not excluded unless we know something about this Deity that would prevent it from creating such universes. An example of such a universe would be Paradise, and it seems unlikely that enthusiasts of the "fine-tuning" argument would be willing to say that the Deity would not create anything like Paradise. But the only way for them to escape from  $P(F|\sim N\&L)\ll 1$  would be for them to assert that the Deity would only, or mostly, create universes that, if they contain life, are  $F$ , and we see no justification for such an assumption.

Kwon makes some other incorrect statements later in his web article. He says that our argument "incorrectly attributes significance to  $P(N|L)$ ." Kwon here appears to have missed the fact that we are talking about Bayesian probabilities. The probability  $P(N|L)$  refers to our universe, and is Your Bayesian prior probability that  $N$  is true,

given that You know that L is true (which must be the case since it is a condition of reasoning that You be alive), but before You learn that F is true. It is a reflection of Your epistemological condition or state of knowledge at a particular moment in time. Thus,  $P(N|L)$  has a perfectly definite meaning in our universe, although the value of  $P(N|L)$  will differ from individual to individual because every individual has different background information (not explicitly called out here but mentioned in our article).

Furthermore, Kwon is incorrect when he states that " $P(N|L)$  is irrelevant to our universe for the same reason that  $P(N|F)$  is irrelevant." We never said that  $P(N|F)$  is irrelevant, only that it is irrelevant for inference. The reason why  $P(N|F)$  is irrelevant for inference is that no sentient being is unaware of L as background information. Every sentient being knows that he is alive and therefore knows that L is true; thus every final probability statement that he makes must be conditioned on L. This is not true of F. There are sentient beings in our universe, indeed in our world, that do not yet know that F is true. Most schoolchildren do not know that F is true, although they know that L is true. Probably most adults do not know that F is true. Thus, Kwon errs in drawing a parallel between  $P(N|L)$  and  $P(N|F)$ .

Kwon started with the perfectly reasonable proposal that "fine tuning" is best defined by  $P(F|N) \ll 1$ , and attempted to derive his result. That he was unable to do this comes as no surprise to us, because one of us [whj] spent the better part of a year trying to get useful information from propositions such as  $P(F|N) \ll 1$ , without success. All such attempts were fruitless, and the reason why is seen in our discussion. For example, suppose we were to assume in addition that  $P(F|\sim N)=1$ . Even then, no useful result can be derived, for from this we can only determine the obvious fact that  $P(F \& L \& \sim N) \ll 1$ , which gives no useful information about the crucial ratio R. The inequality goes in the wrong direction! Thus, "fine tuning"-- $P(F|N) \ll 1$ --tells us nothing useful, which is why in our article we concentrated instead on finding out what "life friendliness"--F--and the WAP can tell us.

Kwon says, "We have always known that F is true for our universe..." This is false. In fact, the suspicion that F is true is relatively recent, going only back to Brandon Carter's seminal papers in the mid-1970's. Earlier, physicists such as Dirac had in fact speculated that the values of some fundamental physical constants (e.g., the fine structure constant) might have been very different in the past, which would violate F, and somewhat later other scientists (for example Fred Hoyle in the early 1950s) have used the assumption that F is true in order to predict certain physical phenomena, which were later found to be the case. Had those observations NOT been found to be true, F would have been refuted, and we would seriously have to consider  $\sim N$ . Even today we do not *know* that our universe is F--"life-friendly"--in the sense that we use the term in our article. We strongly suspect that it is true, but it is conceivable that someone will make a WAP prediction that will turn out to be false and which might refute F.

Kwon incorrectly asserts that the idea that there may be other universes is "simply unscientific." Certainly many highly respected cosmologists and physicists like Andrei Linde (Stanford), Lee Smolin (Harvard) and Alexander Vilenkin (Tufts) and Nobel laureate Stephen Weinberg (Texas) would disagree with this statement. Kwon claims that the hypothesis of other universes "cannot be tested." While we might agree that testing the hypothesis of other universes will be difficult, we do not agree that the hypothesis is untestable, and neither do scientists that work in this area. Some specific tests have been suggested. For example, David Deutsch has proposed specific tests of the Everett-Wheeler interpretation of quantum mechanics commonly known as the "Many-Worlds" hypothesis. And recently an article that proposed another way that other universes might be detected was published (*Science*, Vol. 292, p. 189-190, original paper archived as <http://arXiv.org/abs/hep-th/0103239>). Regardless, our argument is not dependent on the notion that there are many other universes. It stands on its own.

Kwon misunderstands the point of the "god of the gaps" argument. The problem isn't that the gap is being filled by a god, the problem is what happens if the gap is filled by physics. Then the god that filled the gap gets smaller. This is a theological problem, not an epistemological or scientific problem. We agree with Kwon that there are gaps in our physical explanation of the universe that may never be filled; but it is hoping against hope that we will never fill any of the gaps currently being touted by "intelligent design theorists" as proof of supernaturalism. Some of them are certain to be filled in time, and each time this happens, the god of the intelligent designers will be diminished. (In fact, some of them were in fact filled even before the recent crop of "ID theorists" made their arguments--this is true of some of Michael Behe's examples, for which evolutionary pathways had already been proposed even before Behe published his book).

As to Kwon's last point, that we incorrectly claim that "intelligent design theorists" incoherently assert both F and  $\sim F$ . We believe that it is a correct statement that at least some are arguing  $\sim F$ . It is our impression, for

example, that Michael Behe is arguing that it is actually impossible, and not just highly unlikely, for certain "irreducibly complex" (IC) structures to evolve without supernatural intervention, and that is a form of  $\sim F$ . Regardless, even if no one is attempting to argue from  $\sim F$  to  $\sim N$ , our point still stands. Attempts to prove  $\sim N$  that argue from either  $F$  or  $P(F|N) \ll 1$  or both do not work. But attempts to prove  $\sim N$  by showing  $\sim F$  would work. Thus, people making anthropic and "fine tuning" arguments have hold of the wrong end of the stick. They should be trying to show that the universe is not  $F$ . It is clear that showing that the universe is not  $F$  would at one stroke prove  $\sim N$ ; it follows that showing that the universe is  $F$  can only undermine  $\sim N$  and support  $N$ ; this is an elementary result of probability theory, since it is not possible that observations of  $F$  as well as  $\sim F$  would both support  $\sim N$ . Since it is trivially true that observing  $\sim F$  does support  $\sim N$ , observing  $F$  must undermine it. Put another way, it seems to us that Michael Behe--if we understand him--is making the right argument from a logical and inferential point of view, and Hugh Ross is making the wrong argument. If it turns out that Behe is not making the argument we think he is, then it is still the case that Hugh Ross is making the wrong argument.

Kwon makes some remarks about "nontheists" that seem to indicate that he thinks that only "nontheists" would argue as we have. This is not the case. The issue here is whether the "fine tuning" argument is correct. It is exactly analogous to the centuries of work done on Fermat's last theorem. It is likely that most mathematicians thought that the theorem was true for most of that time, yet they continued to reject proofs that had flaws in them. They rejected them not because they thought Fermat's last theorem was false, but because the proofs were wrong. They even rejected Wiles' first attempt at a proof, because it was (slightly) flawed. In the same way a theist can and should reject a flawed "proof" of the existence of God. Our argument is that the fine tuning arguments are wrong, and no one should draw any conclusions about our personal beliefs from the fact that we say that these arguments are wrong.

Conclusion: Kwon's "proof" is fatally flawed. He incorrectly asserts that the only way to keep  $P(N|F\&L)$  from being very small is to assume naturalism a priori. Quite the contrary, the only way to make  $P(N|F\&L)$  small is to assume *supernaturalism* a priori. Kwon apparently does not understand the significance of some of the Bayesian probabilities we use; this is forgivable in a sense since Bayesian probability theory is still misunderstood by most people, even those with some training in probability theory...but it means that Kwon should withdraw these comments until he understands Bayesian probability theory well enough to criticize it. Kwon's assertion that we have always known that our universe is  $F$  is false; his assertion that the existence of other universes is untestable is also false, and in any case is not relevant to our main argument. Finally, he mistakenly thinks that the god-of-the-gap argument somehow tells against science. It does not, since it is purely a theological conundrum, not a scientific one.

Nonetheless, we thank David Kwon for his serious and attentive reading of our article and for his comments. He is the first to attempt a mathematical rather than a polemical refutation of our argument. His argument fails because, as we show here, it isn't possible to derive anything useful from the fine-tuning proposition  $P(F|N) \ll 1$ . When all factors are taken into account, it is clear that the only way to end up with a final result that  $P(N|F\&L) \ll 1$  is to assume at the outset that supernaturalism is almost surely true, thus begging the question.

M. I.  
W. J.  
April 30, 2001

[**Note added 010613:** When we posted this response, we informed Mr. Kwon, so that he could either respond to our criticisms or withdraw his web page. We regret to say that up to now he has done neither.

**Note added 040109:** Kwon has never responded to our criticisms; his web page disappeared when he apparently finished his career as a Berkeley graduate student. It is archived and can be obtained courtesy of the [Internet Wayback Machine](#) via [Makeashorterlink.com](#)]

**Note added 060406:** Another version of Kwon's article appears to have migrated [here](#); We do not know if this site is his or someone else's.

---

## APPENDIX 2: Why one must condition on L

A correspondent who prefers to remain anonymous wrote us as follows (reproduced with permission):

-----Begin Quote-----

Recently I was led to your article with Michael Ikeda called "The Anthropic Principle Does Not Support Supernaturalism,"

<http://quasar.as.utexas.edu/anthropic.html> .

That is quite a striking conclusion.

A key step in your argument, on which you insist repeatedly, is that one must conditionalize on L, the claim that "[t]he universe exists and contains life." The only justification given for this claim, as far as I could find, is that we all know L and we should use everything that we know.

However, this bit of advice leads quickly to a paradox well known to philosophers of science, viz., Clark Glymour's "problem of old evidence."

The problem is that conditionalizing using everything that one knows leads, in some cases, to the absurd conclusion that new theories cannot be confirmed by old evidence. Such a conclusion contradicts common sense and scientific practice. A standard example is the confirmation of Einstein's GR by its entailing the anomalous perihelion precession of Mercury. This precession was known long before Einstein's theory, but Einstein and others have taken it to provide evidence for GR. Surely they were correct. But if one must always use all of the evidence on hand, then Einstein should have reasoned like this:

$E$ =anomalous perihelion precession of Mercury

$T$ =GR

$P(E)=1$  because  $E$  is known.

$P(E|T)=1$  because  $P(E)=1$ .

So Bayes's theorem

$P(T|E) = P(T) P(E|T)/P(E)$  gives  $P(T|E) = P(T)*1/1 = P(T)$ : the probability of GR is not increased by  $E$ ! Some standard responses to this problem involve not using all of one's evidence in some fashion or other.

In short, the only motivation that I find in your paper cited above for conditionalizing on L is one that is widely known among philosophers of science to give absurd conclusions in certain cases. Glymour discusses this problem in "Why I Am Not a Bayesian" in his *Theory and Evidence* (Princeton, 1980), which is also reprinted Curd and Cover, *Philosophy of Science: The Central Issues* (Norton, NY, 1998), with commentary, which is where I am looking at it. A dozen or two responses or counterresponses to the problem can be found in the Philosopher's Index database. Thus a key step in your argument is presently unmotivated in your online paper.

-----End Quote-----

## 2.0 General comments

We have quoted our correspondent's letter in full to address several issues. First, the argument that he attributes to Glymour is wrong. Second, even if it were right, it is not properly applied to the present situation. Third, we will show that for any argument to be *sound*, it *must* include all *background information* which is known to be true and which affects (changes) the likelihood. In the present situation, L has this status. This will motivate in a formal way our assertion that we must condition on L.



Since we have not had an opportunity to read Glymour's original essay, and are therefore not absolutely certain that our correspondent has presented his argument correctly, in the following we will designate the argument our correspondent attributes to Glymour as "Argument A".

## 2.1 Argument A is wrong

We will first deal with Argument A. The argument contains an obvious, fatal flaw.

It is simply *not the case* that the fact that we have observed evidence E entails that  $P(E)=1$ . Since everything in Argument A follows from this mistaken assumption, Argument A is wrong.

$P(E)$  is *not* the probability that E *has been observed*. It is the probability of observing E, *instead of something else*, averaged over *all* theories in the set  $TH = \{T1, T2, \dots\}$  under consideration, with weights proportional to the prior probabilities of the theories in TH. [We assume that every theory T in TH has positive prior probability, i.e.,  $P(T) > 0$  for all T in TH]. E is a candidate from the set of *all possible outcomes*  $EV = \{E1, E2, \dots\}$  that these theories predict *could be observed*. Therefore,  $P(E)$  is in general *not* equal to 1, even *after* you have observed E. Indeed,  $P(E)$  is the same number before you observe evidence E, after you observe evidence E, or even if you never observe evidence E. It is equal to 1 if and only if *every* theory in TH predicts that *only* E could *ever* be observed.

As [Tom Loredo](#) pointed out to us when we showed him Argument A, "Time plays the same role in probability theory as it does in logic, i.e., no role whatsoever." This means the probability calculus, like the logic calculus, produces sound results, independently of when you learn the truth or falsity of any of the premises in the statement. This fact becomes obvious when one learns that in the limit when propositions are definitely true or false, probability theory reduces to ordinary logic, as a consequence of a theorem due to [Cox](#) (1946). For a transparent discussion of this relationship, see pp. 12-23 of the following [lecture](#) by Tom Loredo.

$P(E)$  is known technically as the *marginal likelihood*, and it is *correctly* computed using a specific formula involving another quantity known as the *likelihood function*. It is *never* computed from a naive statement such as "I've observed E, therefore  $P(E)=1$ ." In what follows we will define these quantities and show how Argument A should have calculated  $P(T|E)$  from  $P(T)$  and knowledge of E. We will also show precisely where Argument A went wrong.

### 2.1.1 Sampling distribution, likelihood, and marginal likelihood

In Bayesian inference, one is interested in learning how the inclusion of evidence E *changes* our belief about the plausibility of various theories, compared to what one believed about those theories *without* that evidence. This means that one should start with  $P(T)$ , *unconditioned* on E (i.e., without that evidence), and given E, calculate  $P(T|E)$  (with that evidence). This is what Argument A alleges to do, but does incorrectly. For clarity, we will restrict ourselves to just two theories,  $\{T1, T2\}$ .

Standard Bayesian theory starts with  $P(E|T)$ . This is generally not equal to 1, *even if we have already observed evidence E*. Technically, when  $P(E|T)$  is conditioned on a fixed theory T and considered as a function of the various E in EV, it is known as the *sampling distribution under T*. It tells us, on the assumption that T is true, the probability of observing each outcome E, where E ranges over all the possible outcomes in EV. Since it is a probability (when considered as a function of E), its sum over all the possible values of E is 1:

$$P(E1|T) + P(E2|T) + P(E3|T) + \dots = 1$$

Because of this equation,  $P(E|T)$  can be equal to 1 *only* when the theory T *predicts* that it is *impossible* to observe any outcome other than E. This is true regardless of whether E has already been observed, is yet to be observed, or even if it is never observed.

The sampling distribution (that is, the *function*  $P(E|T)$ ) doesn't care what evidence we actually observe. It is constructed *independently* of any observed evidence, and has the *same* numerical value for each of its arguments *after* evidence E is observed as it had before. It is therefore only a tool to describe a particular theory T, and *not* a description of evidence that may or may not have been observed.

In Bayesian inference, one is interested in comparing several theories. For each theory  $T$  in  $TH$ , we construct its sampling distribution  $P(E|T)$ , which tells us how likely it is, under each theory, that we would observe evidence  $E$  (ranging over all the alternatives contained in  $EV$ ). Once we observe a *particular* piece of evidence  $E$ , we are able to consider  $P(E|T)$  as a function of the second argument  $T$ . The function of  $T$  that we get by fixing  $E$  at its observed value and allowing  $T$  to vary over all theories in  $TH$  is known as the *likelihood function*. It is *not* a probability, and it is *not* normalized (the sum of  $P(E|T)$  over all  $T$  doesn't have to add up to 1). It can even be multiplied by an arbitrary positive constant  $C$  (independent of  $T$ ) without affecting any inferences.

In the general relativity example, we are interested in comparing theory  $T1$  (say general relativity) with theory  $T2$  (say Newtonian physics). The likelihood function is given by the values of  $P(E|T1)$  and  $P(E|T2)$ , evaluated with the actual evidence  $E$  we have observed. Suppose there are only two possible outcomes of our experiment,  $E1$ ="observe anomalous perihelion precession of Mercury" and  $E2$ ="observe no anomalous perihelion precession of Mercury".

The sampling distribution under the two theories is as follows:

$$P(E1|T1)=1, P(E2|T1)=0 \\ P(E1|T2)=0, P(E2|T2)=1$$

This is because  $T1$  predicts that we *must* observe anomalous perihelion motion, and  $T2$  predicts that we *cannot* observe anomalous perihelion motion[1]. It doesn't matter when  $E1$  or  $E2$  is observed, these probabilities are dictated by the theory alone, and not by any observations that might or might not have been made. Historically,  $E1$  was observed almost a century before general relativity was proposed. But even so, the sampling distributions under each theory, which are *always* constructed *independently* of any evidence, describe only what the *theories* say we can observe, and are as given above.

Once we say to ourselves, "We observed  $E1$ , not  $E2$ ", we can refine the situation. For now we can write down the *likelihood function*, which is a function of the *second* argument, with the first argument *fixed* at the observed  $E1$ . Consulting the above four equations, we find that the likelihood is given by

$$P(E1|T1)=1, P(E1|T2)=0$$

Note: Even though we now know that  $E1$  is true,  $P(E1|T2)$  does not suddenly change its value to 1 as Argument A would seem to say, but (in this example) remains equal to 0. To repeat what we've said before, this is because for every theory  $T$ , the *function*  $P(E|T)$  describes the theory  $T$ , independently of any evidence  $E$  we may have actually observed.

Next, we must assign priors to  $T1$  and  $T2$ . As an illustration, set  $P(T1)=P(T2)=1/2$ . With this assignment, we can compute the *marginal likelihood*,  $P(E1)$ . This is *always* computed by expanding  $P(E1)$  as follows:

$$P(E1)=P(E1|T1)P(T1)+P(E1|T2)P(T2)=1*1/2+0*1/2=1/2$$

Note: Argument A claims that  $P(E1)=1$ ; this is manifestly false.  $P(E1)$  is just a normalization constant, designed to guarantee that the posterior probability is a normalized probability on the theories  $T1, T2, T3, \dots$ . Thus,  $P(T1|E1)+P(T2|E1)+\dots=1$ . Routine calculation shows that this requires us to set

$$P(E1)=P(E1|T1)P(T1)+P(E1|T2)P(T2)$$

Finally, we calculate the posterior probability of  $T1$ , given  $E1$ , this time correctly:

$$P(T1|E1)=P(E1|T1)P(T1)/P(E1)=1*1/2/(1/2)=1$$

Notice that the calculation results in a posterior probability  $P(T1|E1)$  that is *different* from the prior probability  $P(T1)$ ! Contrary to Argument A's assertion, we *can* learn from old data, and the inclusion of old evidence  $E1$  *does* support  $T1$  by showing (in this case) that  $P(T1|E1)>P(T1)$ .

## 2.1.2 What went wrong?

Evidently, something has gone wrong. A clue as to what is wrong with Argument A can be gleaned from its (incorrect) claim that  $P(E)=1$ . Evidently, the thinking is: E is old evidence, I know that E is true, therefore  $P(E)=1$ . This reasoning is incorrect, because the *only* correct way to calculate  $P(E)$  is through the expression we have displayed above. Nonetheless, from this insight into the thinking, we can infer what's gone wrong. Argument A is actually conditioning on the fact that E has already been observed, without displaying that conditioning explicitly. Thus, what Argument A calls  $P(E)$  is actually  $P(E|E)$ , which *is* equal to 1. It regards E as already-known *background information*.

Bayes' theorem, written with background information B, takes the form

$$P(T|E,B)=P(E|T,B)P(T|B)/P(E|B)$$

If E is regarded as background information B, simple substitution yields

$$\begin{aligned} P(T|E) &= P(T|E,E) = P(E|T,E)P(T|E)/P(E|E) \\ &= P(T|E), \end{aligned}$$

since trivially  $P(E|E)=P(E|T,E)=1$ . This statement correctly demonstrates that if we start with  $P(T|E)$  as the prior on T, then inserting E into Bayes' theorem as evidence does not change anything. The posterior equals the prior. Bayes' theorem does not allow you to use the same evidence twice.

But the rub is that the *real* prior  $P(T)$  has *never* used evidence E, not even once. Argument A is claiming that if evidence is old, Bayes' theorem shows that  $P(T|E)=P(T)$ . But that is false. If one substitutes  $P(T)$  for  $P(T|B)$  on the right hand side of Bayes' theorem above, one gets the "equation"

$$P(T|E,B)=P(E|T,B)P(T)/P(E|B) \text{ (???)},$$

which is *not* a theorem and is in general false. If we were to set  $B=E$  in this expression, we would get  $P(T|E)=P(T)$ , but since the expression is *not a theorem*, the argument is *invalid*.

The late E. T. Jaynes, in his book *Probability Theory: The Logic of Science* (Cambridge University Press), put his finger on the problem when he pointed out that *failure to condition properly on all known and relevant background information* often leads to *apparent* paradoxes in probability theory. These apparent paradoxes disappear when the correct conditioning is displayed explicitly, as we have done above.

The attentive reader will also notice that Jaynes' dictum to condition on all known and relevant background information is *precisely* what we have been saying all along in our discussion of the anthropic principle. L is known true *a priori*, and affects the likelihood, therefore one must condition on L in order to avoid apparent "paradoxes" such as Argument A.

## 2.2 Argument A is misapplied here

Even if Argument A were correct, it is irrelevant to our discussion. The reason is simple. Our interest is in what happens when *new* information F is presented to someone who *already knows that L is true*, and who has *evaluated his priors in the light of the fact that L is true*. In other words, we are only interested in what happens when a Bayesian calculating machine that knows that L is true is given, for the first time, the *new* information that F is true. As we point out in our article, *every sentient being* knows from the first time that it becomes sentient that L is true. But F is genuinely new information, only known to some physicists since c. 1950 at the earliest, and still unknown to the majority of human beings. The "fine tuning" argument isn't "What do you think about God, when you learn that you are alive?" but "What do you think about God, when you learn that the universe is (apparently) fine-tuned or life-friendly?"

This means that the argument about "old evidence" is not even relevant to our discussion, since we are talking about what happens when you learn that F is true, not about what happens when you learn that L is true (which you already knew...). The "old information" is L, but the "new information" is F.

### 2.3 Motivation for conditioning on L

Having dealt with Argument A, we now deal with the objection that our requirement to condition on L is unmotivated. We motivate the conditioning on L by appealing to the principle that arguments should be sound.

For an argument to be *sound*, it must be both *factually correct* and *valid*.

*Factually correct* means that all its premises are *true*. *Valid* means that the conclusions follow from the premises.

For example, consider the argument:

All men are immortal  
Socrates is a man  
Therefore, Socrates is immortal

This argument is *valid*, because the conclusion logically follows from the premises. However, the premise "All men are immortal" is *factually incorrect*, therefore the argument is *unsound*.

Conversely, the argument

All men are mortal  
Socrates is mortal  
Therefore, Socrates is a man

is *unsound* because it is invalid, even though it is factually correct. The premises are true, but the conclusion does not follow from the premises.

Similarly, a Bayesian calculation is *valid* if it uses the probability calculus correctly, and *factually correct* if all of its premises (assumptions) are correct. It is *sound* if it is both *valid* and *factually correct*.

We will show that if one attempts to ignore true background information B in the likelihood function  $P(E|T, B)$ , and if B actually affects the values taken on by the likelihood, the argument will *not* be factually correct, and therefore the argument will be *unsound*.

Suppose that I claim to draw a conclusion about T from evidence E, and claim that only  $P(E|T)$ , unconditioned on B, needs to be considered as the likelihood.

You are skeptical of this. You note that, regardless of what B is, one can always write

$$P(E|T) = P(E|B, T)P(B|T) + P(E|\sim B, T)P(\sim B|T)$$

You also note that, by Bayes' theorem,

$$P(B|T) = P(T|B)P(B)/P(T) \text{ and } P(\sim B|T) = P(T|\sim B)P(\sim B)/P(T)$$

Plugging these expressions into the previous we find

$$P(E|T) = \frac{P(E|B, T)P(T|B)P(B) + P(E|\sim B, T)P(T|\sim B)P(\sim B)}{P(T|B)P(B) + P(T|\sim B)P(\sim B)},$$

where the denominator  $P(T)$  has been expanded using the formula we explained above.

You challenge me to tell you whether B is true or not. If I know that B is true, regardless of how I know it, I am obliged to tell you the truth. If I fib to you, then the argument I am trying to make will automatically be *factually incorrect*, since some premises will be false, and hence my argument will be *unsound*.

Thus, I am obliged to report to you that B is true, so that  $P(\sim B)=0$ . You will then calculate

$$P(E|T)=P(E|B,T) \text{ for all values of E and T}$$

You will conclude that you cannot leave B out of the conditioning on the likelihood. My attempt to avoid conditioning on B has failed: If the presence of B affects the likelihood  $P(E|B,T)$ , we must use a  $P(E|T)$  that reflects that information by being numerically equal to  $P(E|B,T)$  for all values of E and T. Thus, the actual likelihood is  $P(E|B,T)$ , despite my attempt to pull the wool over your eyes by not mentioning B when I wrote down the likelihood. Only if E is independent of B is it justified to use just  $P(E|T)$ , because independence means that  $P(E|B,T)=P(E|T)$ .

Jaynes' dictum, "Condition on everything you knew before the new evidence," is validated.

Specifically, in the example at hand,  $E=F$ ,  $T=N$  and  $B=L$ . So, we have shown that even if I attempt to leave L out of the equation, we find that numerically, for all values of F and N,

$$P(F|N)=P(F|L,N)$$

Specifically, we know that the *sampling distribution* of F under L and N is 1, so  $P(F|L,N)=1$ ; for if we have a naturalistic universe that contains life, this entails that F is true. And, we know that the sampling distribution of F under L and  $\sim N$  is  $\leq 1$ , since we cannot logically rule out non-naturalistic universes with life that are  $\sim F$ .

Therefore, we compute that the Bayes factor  $P(F|L,N)/P(F|L,\sim N) \geq 1$ , i.e., observing that F is true supports (or at least does not undermine) our belief in N. This is precisely the same conclusion we obtained before. Even if I try to pull the wool over your eyes by failing to mention L in the conditioning, the above argument shows that  $P(F|N)/P(F|\sim N) \geq 1$  when the correct likelihood is used.

----

[1] This is only a very good approximation. The value of the anomalous GR precession is about 43"/century, close to the observed value. But actually, if through extraordinary bad luck the observational errors just happened to be horrible, it might be possible to have observed an anomalous 43"/century even if the true value were zero, and vice versa. Strictly speaking, therefore, the probabilities in the table should be very close to 1 or 0, but ought to differ from these numbers by a very small quantity.

----

M.I.  
W.J.  
April, 2006