Absence of Evidence and Evidence of Absence: Evidential Transitivity in connection with Fossils, Fishing, Fine-Tuning, and Firing Squads

Elliott Sober
University of Wisconsin, Madison

5185 Helen C. White Hall
Philosophy Department
University of Wisconsin, Madison
Madison, WI 53706
Tel 608 263 3700 ersober@wisc.edu Fax 608 265 3701

Abstract: "Absence of evidence isn't evidence of absence" is a slogan that is popular among scientists and nonscientists alike. This paper assesses its truth by using a probabilistic tool, the Law of Likelihood. Qualitative questions ("Is *E* evidence about *H*?") and quantitative questions ("How much evidence does *E* provide about *H*?") are both considered. The paper discusses the example of fossil intermediates. If finding a fossil that is phenotypically intermediate between two extant species provides evidence that those species have a common ancestor, does failing to find such a fossil constitute evidence that there was no common ancestor? Or should the failure merely be chalked up to the imperfection of the fossil record? The transitivity of the evidence relation in simple causal chains provides a broader context, which leads to discussion of the fine-tuning argument, the anthropic principle, and observation selection effects.

Keywords: anthropic principle, Bayesianism, common ancestry, evidence, fine-tuning, fossils, likelihood.

Absence of Evidence and Evidence of Absence: Evidential Transitivity in connection with Fossils, Fishing, Fine-Tuning, and Firing Squads

Elliott Sober

University of Wisconsin, Madison

1. Introduction

Scientists often say that *absence of evidence isn't evidence of absence*. In fact, they don't just *say* this; they sometimes *wear* it too. The American Statistical Association is selling a T-shirt on which the motto is proudly displayed.[1] This epistemological aphorism is just the sort of thing that philosophers love to analyze. The tools of probability theory have been brought to bear on other principles about evidence -- that a varied body of evidence provides stronger support for a hypothesis than evidence all of the same kind, that a theory that unifies a body of evidence is better supported by that evidence than is a theory that provides separate explanations for separate parts of the evidence, and so on. But, as far as I know, probabilists have not had a go at this one.

There has been some philosophical work on the motto that doesn't use a probability framework and it provides a good way of isolating the problem I want to address. Here are two example arguments from Douglas Walton's insightful 1996 book, *Arguments from Ignorance*:

I do not have any evidence that it is raining here and now.

It is not raining here and now.

I do not have any evidence that there is a storm on the surface of Jupiter now.

There is no storm on the surface of Jupiter now.

Though neither argument is deductively valid, it is easy to see how each can be turned into a valid argument by adding a premise. The arguments have the form:

[1] Go to http://www.amstat.org/ASAStore/Absence_of_Evidence_Adult_T-sh_P108.cfm

I do not have any evidence that *p* is true.

---

p is false.

Just add the premise

(P_1) If *p* were true, then I'd have evidence that *p* is true.[2]

This further premise may be true in the case of the rain. Suppose, as in Walton's example, that I am sitting in a house with a tin roof and that I'd hear the characteristic pitter-patter if rain were falling. It is easy to imagine that the extra premise P_1 is false in the case of the storm on Jupiter; suppose, instead, that

(P_2) If *p* were true, then I'd have no evidence that *p* is true.

The Jupiter example is enough to show that the motto "absence of evidence isn't evidence of absence" is sometimes true and the rain example is enough to show that it is sometimes false.[3] This is because P_2 is true of some propositions in some circumstances and the same goes for P_1. So let us agree that absence of evidence does not *logically entail* that you have evidence of absence. And let us also agree that there are situations in which absence of evidence *is* evidence of absence. What more is there to say about the motto than this?

My goal in this paper is to explore cases in which P_1 and P_2 are *both* false. Such cases arise when having evidence is a matter of chance. To see why chance can render both P_1 and P_2 false, consider an analogy. Suppose, when you toss a coin, that the probabilities of the two outcomes are each strictly between 0 and 1. If so, both the following conditionals are false:

If you tossed the coin, it would land heads.

---

[2] Walton (1996, p. 252) says that *modus tollens* is the form of a basic type of *ad ignorantiam* argument.

[3] Chris Stephens has drawn my attention to two other examples. The administration of George W. Bush justified its 2003 invasion of Iraq by saying that there was evidence that Iraq possessed "weapons of mass destruction." After the invasion, when none turned up, Donald Rumsfeld, who then was Bush's Secretary of Defense, addressed the doubters by invoking the motto; see http://en.wikiquote.org/wiki/Donald_Rumsfeld. Carl Sagan (1997) does the same when he considers the fact that we have not yet found evidence that extra-terrestrial intelligence exists.

If you tossed the coin, it would land tails.

In considering cases in which a probabilistic process governs the gaining of evidence, I'm interested in seeing, not just when the motto is *exactly* true, but also when the motto is *close* to being true. Even when absence of evidence provides *some* evidence of absence, when is the evidence *substantial* and when is it *negligible*?

In the next section, I discuss the epistemological framework I'll use to analyze these questions about absence of evidence. In the section after that, I consider the bearing of fossils on hypotheses of common ancestry. According to evolutionary biologists, observing that a fossil has phenotypic features that are intermediate between the features of two extant species is evidence that those species have a common ancestor. Creationists do not always agree, but they often maintain that *failing* to find a fossil intermediate is evidence *against* the hypothesis of common ancestry. Evolutionary biologists usually reply to this creationist assertion with our slogan – that absence of evidence for a common ancestor isn't evidence that there was no such thing. Analyzing this problem leads to some general conclusions about the epistemological features of a causal chain that runs from some state of the world, to a trace, and then to an observation of that trace's properties. This in turn leads to issues concerning the fine-tuning argument, the anthropic principle, and observation selection effects.

2. The Law of Likelihood

The epistemology that I'll develop in this paper will be based on the Law of Likelihood (Hacking 1965, Edward 1973, Royall 1997):

> *The Law of Likelihood.* Evidence E favors hypothesis H_1 over hypothesis H_2 precisely when $Pr(E \mid H_1) > Pr(E \mid H_2)$. And the degree to which E favors H_1 over H_2 is measured by the likelihood ratio $Pr(E \mid H_1)/Pr(E \mid H_2)$.

Notice that this principle has two parts. The first is qualitative while the second is quantitative. I won't attempt to provide a full defense of the Law, but I do want to make a few comments before I put it to work (for more details, see Sober 2008).

Let's begin with the qualitative part of the principle. This is something that Bayesians should embrace. For Bayesians, observational evidence can modify one's degrees of belief in various hypotheses only by way of likelihoods. The odds version of Bayes' Theorem shows why:

$$\frac{\Pr(H_1 \mid E)}{\Pr(H_2 \mid E)} = \frac{Pr(E \mid H_1)}{Pr(E \mid H_2)} \times \frac{Pr(H_1)}{Pr(H_2)}.$$

$$\text{ratio of posteriors} = \text{ratio of likelihoods} \times \text{ratio of priors}$$

The ratio of the posterior probabilities can differ from the ratio of priors only if the likelihoods are different. And the more the likelihood ratio deviates from unity, the more the ratio of posterior probabilities will differ from the ratio of priors. Another reason why Bayesians should smile on the qualitative part of the Law of Likelihood is that it fits in with the Bayesian theory of confirmation:

> *The Bayesian Theory of Confirmation*: Evidence *E* confirms hypothesis *H* if and only if $Pr(H \mid E) > Pr(H)$.

This view of confirmation is equivalent with the following

> Evidence *E* confirms hypothesis *H* if and only if $Pr(E \mid H) > Pr(E \mid notH)$.

For Bayesians, when a hypothesis is confirmed by evidence, the evidence favors that hypothesis over its own negation, where "favoring" means what the Law of Likelihood says it means.

Although Bayesians are (or should be) friends of the (qualitative) Law of Likelihood, you don't have to be a Bayesian to find merit in the Law. Likelihoodists hold that the principle describes how evidence should be interpreted even when they refuse to assign prior or posterior probabilities to the hypotheses under test because these quantities are insufficiently "objective." They also tend to reject the Bayesian theory of confirmation because they do not think that testing a theory must always pit the theory against its own negation. For example, they grant that it makes sense to talk about the probability of Eddington's eclipse data, conditional on the general theory of relativity's being true, but decline to talk about the probability of the data, conditional on the "catch-all hypothesis" that is the GTR's negation (Earman 1992). What likelihoodists think is important is the testing of specific theories against each other; this is what Eddington did when he tested the GTR by comparing its predictions with those that issue from Newtonian theory.

I now turn to the second part of the Law of Likelihood, which says that strength of evidence should be measured by the likelihood ratio. Although this measure is controversial among Bayesians (Fitelson 1999), it has not been discussed much by those who embrace the qualitative part of the Law of Likelihood. It should be. Likelihoodists want a measure that does not depend on prior or posterior probabilities, and this knocks out a lot of candidates. For example, the following candidate is out the window:

The difference measure: The degree to which E favors H over $notH = Pr(H \mid E) - Pr(H)$.

Likelihoodists want a measure of evidential favoring whose value depends only on the likelihoods of the two hypotheses. But why opt for the likelihood *ratio*? For example, why not use the likelihood *difference*? One reason is suggested by a pattern that arises when there are multiple pieces of evidence that are independent of each other, conditional on each of the hypotheses considered. Suppose, for example, that

$\Pr(E_i \mid H_1) = 0.99$, for each of the observations $E_1, \ldots, E_{1000}$.
$\Pr(E_i \mid H_2) = 0.3$, for each of the observations $E_1, \ldots, E_{1000}$.

With conditional independence, we have

$$\Pr(E_1 \,\&\, \ldots \&\, E_{1000} \mid H_1) = (0.99)^{1000} \text{ and } \Pr(E_1 \,\&\ldots\&\, E_{1000} \mid H_2) = (0.3)^{1000}.$$

The likelihood of each of these hypotheses, relative to the thousand observations, is very close to zero, so their difference is tiny; however, the ratio of the two likelihoods is $(33)^{1000}$, which is huge. Since each of these thousand (conditionally independent) observations favors H_1 over H_2, the thousand observations should do so more powerfully than any of them does singly. This recommends the likelihood ratio over the likelihood difference as a measure of strength of evidence.

Figure 1. Smith and Jones differ in their inclinations to place different orders for breakfast. The breakfast order of *toast and eggs* provides evidence about which of them placed the order, although the fact that the order included *toast* does not, and neither does the fact that the order included *eggs*.

Pr(-- | Smith)

		Eggs +	Eggs -
Toast	+	0.4	0.1
	-	0.1	0.4

Pr(-- | Jones)

		Eggs +	Eggs -
Toast	+	0.1	0.4
	-	0.4	0.1

The Law of Likelihood reflects the fact that the evidence relation can be sensitive to logically strengthening and weakening one's description of the evidence. Suppose that proposition S is logically stronger than proposition W, meaning that S entails W, but not conversely. It can turn out that S favors H_1 over H_2 while W fails to do so, and it also is possible for W to favor H_1 over H_2 even though S fails to do so. Here are two examples that illustrate these points. Suppose you are a cook in a restaurant. The waiter brings an order into the kitchen – someone in the dining room has ordered toast and eggs for breakfast. Does this evidence discriminate between the hypothesis that your friend Smith placed the order and the hypothesis that your friend Jones did so? You know the eating habits of each; the probabilities of different breakfast orders, conditional on Smith's placing the order, and conditional on Jones' placing the order, are shown in Figure 1. These probabilities give rise to the following curious fact: the order's being for *toast and eggs* favors Smith over Jones (since 0.4 > 0.1); but the fact that the customer asked for *toast* provides no evidence on this question (since 0.5 = 0.5) and the fact that the customer asked for *eggs* doesn't either (since, again, 0.5 = 0.5). Here the whole of the evidence is more than the sum of its parts.

Figure 2 depicts the opposite pattern in which a new set of inclinations is attributed to your friends. If Smith and Jones are disposed to behave as described, an order of *toast and eggs* fails to discriminate between the two hypotheses (since 0.4 = 0.4). However, the fact that the order included *toast* favors Smith over Jones (since 0.7 > 0.6) and the same is true of the fact that the order included *eggs* (since 0.6 > 0.4). Here the whole of the evidence is *less* than the sum of its parts.

Figure 2. A new set of breakfast inclinations for Smith and Jones. Now the breakfast order of *toast and eggs* provides no evidence about which of them placed the order, though each part of the order favors Smith over Jones.

Pr(-- ∣ Smith)

Toast \ Eggs	+	-
+	0.4	0.3
-	0.2	0.1

Pr(-- ∣ Jones)

Toast \ Eggs	+	-
+	0.4	0.2
-	0	0.4

This point about strengthening and weakening the description of the evidence is relevant to assessing the motto about absence of evidence because the proposition

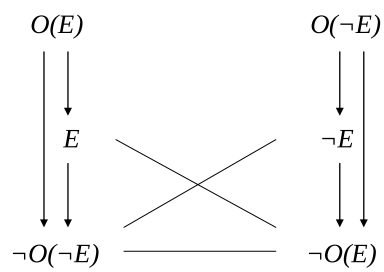
I observe that *E* is true

is logically stronger than the proposition *E*, which in turn is stronger than

I do not observe that *E* is false.

The relevant logical relationships are shown in Figure 3. I'm assuming here that the locution "*S* observes that *E*" is *factive*; if you observe that *E* is true, this entails that *E* is true. Of course, it is possible to *think* you've observed that *E* is true even though, as it happens, *E* is false, but that's different. Suppose that *E* would be evidence that some entity exists (or that some process occurs); *E* is "evidence of presence" and therefore ¬*E* would be evidence of absence. Failing to observe that *E* is true *(¬O(E))* is a case of absence of evidence. In shifting from ¬*E* to ¬*O(E),* one moves from evidence of absence to absence of evidence. This shift from a logically stronger to a logically weaker proposition can engender a change in evidential import, as the stories about toast and eggs suggest. Understanding when this occurs, and how much difference it makes when it does occur, are key to assessing what truth there is in the motto.

Figure 3. Logical relationships among six propositions. $O(\varphi)$ means you observe that φ is true. Arrows represent entailment and lines represent compatibility without entailment. Propositions not joined by an arrow or a line are incompatible.



3. Intermediate Fossils as Evidence for Common Ancestry

With this stage-setting in place, I now turn to an example from evolutionary biology. Suppose you are wondering whether two species that you now observe, *X* and *Y*, have a common ancestor. To bring evidence to bear on this question, you might look at the similarities and differences (both phenotypic and genetic) that characterize the two species. But the traits of a third object might be relevant as well. Suppose you observe that there is a fossil whose trait values are intermediate between those exhibited by *X* and *Y*. How does the discovery of this fossil intermediate affect the question of whether *X* and *Y* have a common ancestor?

Creationists frequently claim that the absence of intermediate fossil forms is evidence against common ancestry. Evolutionists reply by pointing to the numerous intermediate fossils that have been discovered; these link dinosaurs with birds, land tetrapods with fish, reptiles with mammals, and land mammals with whales. Of course, if you discover an *I* that is intermediate between species *X* and *Y*, the question can still be raised as to where the forms are that fall between *X* and *I* and between *I* and *Y*. There will always be "gaps;" they just get narrower.

Biologists do not interpret these gaps, whether they are narrow or wide, as evidence against common ancestry. Gaps are simply chalked up to "the imperfection of the fossil record;" fossils often don't get formed and even when they do, it is easy enough for them to be destroyed or for biologists to fail to find them. When evolutionists reply in this way, are they guilty of wanting to have their cake and eat it too? If finding intermediate fossils is evidence *for* common ancestry, isn't failing to find them evidence *against*?

Figure 4. If the evolutionary process is gradual, the *CA* hypothesis predicts the existence of organisms that had intermediate forms, regardless of the character state of the common ancestor *Z*. What the *SA* hypothesis predicts depends on the states of the postulated ancestors Z_1 and Z_2.

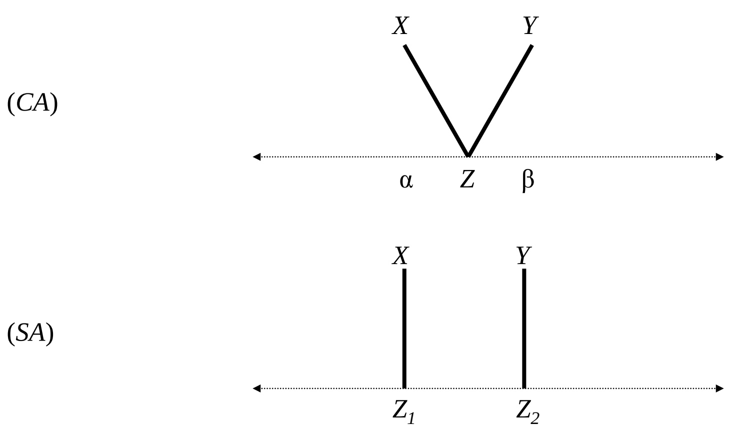


Figure 4 depicts what the hypothesis of common ancestry (*CA*) and the hypothesis of separate ancestry (*SA*) say about the existence of intermediate forms. We observe the character states of the extant species *X* and *Y* (α and β, respectively). If evolutionary change proceeds gradually, there *must* be intermediate forms if *X* and *Y* have a common ancestor (*Z*). Slide *Z* along the scale that represents its possible trait values; no matter what character state *Z* occupies, the lineage leading from *Z* to *X* or the lineage leading from *Z* to *Y* must contain organisms whose trait values fall between the values α and β. This is not true for the separate ancestry hypothesis. If the lineage passing through Z_1 on its way to *X* never has a trait value that is greater than α and the lineage passing through Z_2 on its way to *Y* never has a trait value that is less than β, then there are no ancestors in the lineages leading to *X* and *Y* that have trait values that fall between α and β.

Figure 5. Either *X* and *Y* have a common ancestor (*CA*) or they do not (*SA*). Cells represent probabilities of the form Pr(± intermediate │ ±*CA*). Gradualism is assumed.		
	CA	***SA***
There existed an intermediate.	1	*q*
There did not.	0	*1-q*

Historically, the hypothesis of separate ancestry has often been associated with the claim of *evolutionary stasis* – the thesis that ancestors were, in the main, just like their descendants.[4] However, the *SA* hypothesis is not logically committed to stasis. We can and should separate the *SA* hypothesis from the assumption that lineages do not change trait values.[5] The upshot is that *CA* and *SA* provide different answers to the question of whether intermediate forms once existed; given the assumption of gradualism, *CA* answers that they *must have existed* while *SA*'s reply is that they *may have*. It is only a short step to the following likelihood inequality:

Pr(an organism intermediate between *X* and *Y* existed │ *CA*) >
Pr(an organism intermediate between *X* and *Y* existed │ *SA*).

Given gradualism, the first of these likelihoods has a value of unity. If the *SA* hypothesis allows that there is some chance that the lineages leading to *X* and *Y* never strayed into the "intermediate zone" between α and β, then the second likelihood (which I will call "*q*") is less than unity. If we add to the hypothesis of separate ancestry the stronger assumption of evolutionary stasis, the second likelihood has a value of zero. These points are summarized in Figure 5. Notice that entries in each column must sum to unity. If we use the likelihood ratio to represent how strong the evidence is that favors one hypothesis over the other, we obtain an asymmetry. If there is an intermediate form, this favors *CA* over *SA*, and the strength of this favoring is represented by the ratio $1/q$. This ratio has a value greater than unity if $q<1$. On the other hand, if there is no intermediate, this *infinitely* favors *SA* over *CA*, since $(1-q)/0 = \infty$ (again assuming that $q<1$).[6] The non-existence of an intermediate form would have a far more profound evidential impact than the existence of an intermediate.

[4] Lamarck is an exception to this pattern; he held, for example, that current human beings and current dogs don't have a common ancestor, though each line has evolved (or will evolve) through the same preordained sequence of stages. Our lineage is older since we are more complex.
[5] The same point holds for the historical association of separate ancestry and intelligent design.
[6] Since dividing by zero is not defined, perhaps the point is put better by saying that the nonexistence of an intermediate organism would *refute* CA (assuming gradualism) and would thereby provide the strongest possible discrimination between the two hypotheses.

Although the assumption of gradualism plays a role in these likelihood comparisons, it is important to remember that gradualism is not plausible for some traits. Consider the example of chromosome number. There is no iron law of evolution that says that a lineage that evolves from 12 pairs of chromosomes to 24 must evolve from 12 to 13 to 14 to … 23 to 24. Polyploidy (the doubling or tripling of chromosome number) is a known process.[7] Still, gradualism is usually assumed when the evolution of a continuous character is discussed, and discussion of "intermediate" forms usually involves continuous characters.

Figure 6. Either *X* and *Y* have a common ancestor (*CA*) or they do not (*SA*). Cells represent probabilities of the form Pr(± we have observed an intermediate | ±*CA*). Gradualism is assumed.

	CA	***SA***
We have observed an intermediate.	*a*	*qa*
We have not.	*1-a*	*1-qa*

We now can turn to the accusation that evolutionists play a game of "heads I win, tails you lose" when they appeal to the imperfection of the fossil record to excuse the fact that no fossil that is intermediate between *X* and *Y* has yet been observed. The key is to not confuse the *existence* of intermediates with our *observing* such intermediates. As we have seen, the hypothesis of common ancestry is committed to the existence of intermediates so long as gradualism is correct. But the hypothesis of common ancestry does not guarantee that we will have *observed* those intermediate forms. That depends on how often they fossilize, on how long those fossils last, and on how much fossil hunting paleobiologists undertake. To model the probability of *observing* an intermediate form, or of failing to do so, conditional on each of the two hypotheses, we need the old quantity *q* = Pr(there exists an intermediate | *SA*), but we'll also use something new:

(SO) *a* = Pr(we have observed an intermediate | *CA* & there exists an intermediate) =
Pr(we have observed an intermediate | *SA* & there exists an intermediate).

This proposition expresses an assumption -- that the probability of observing an intermediate, if one exists, is the same, regardless of whether *CA* or *SA* is true; that is, the existence of an intermediate *screens off* the observation of an intermediate from each of the genealogical

7 Developmental genetics provides numerous examples (e.g., hox genes) in which small genetic changes induce discontinuous phenotypic changes; see Carroll (2005) for an introduction.

hypotheses.[8] Whereas the parameter *q* describes the process by which genealogies give rise to intermediate organisms, the parameter *a* describes the process of *observing* intermediates once they exist. The relevant likelihoods are shown in Figure 6. Notice that the likelihood ratio of *CA* to *SA*, given that we have observed an intermediate fossil, is *1/q*. In this respect, Figures 5 and 6 agree; it makes no difference whether we describe the evidence by saying that an intermediate exists or by saying that we have observed that an intermediate exists. However, when we have *not* observed an intermediate, the likelihood ratio of *SA* to *CA* takes on the value

$$\frac{\Pr(\text{we have not observed an intermediate} \mid SA)}{\Pr(\text{we have not observed an intermediate} \mid CA)} = \frac{1\text{-}qa}{1\text{-}a}.$$

Here the shift from a logically stronger proposition (that no intermediate organism exists) to a logically weaker proposition (that we have not observed an intermediate) makes a difference; instead of considering an infinite likelihood ratio, we now confront one that is finite (if $a<1$). This ratio is greater than unity if $a>0$ and $q<1$. As long as there is *some* chance that we'll observe an intermediate if one exists, and there is *some* chance that intermediates will not exist if the separate ancestry hypothesis is true, the failure to observe a fossil intermediate favors *SA* over *CA*. In this broad circumstance, absence of evidence (for a common ancestor) *is* evidence that there was no such thing. The motto – that absence of evidence isn't evidence of absence -- is wrong.

Thus far we have a symmetry -- observing an intermediate favors *CA* over *SA* and failing to so observe has the opposite significance (provided that *a* and *q* are constrained as just described). This qualitative symmetry leaves open whether there is a quantitative asymmetry. Maybe observing an intermediate favors *CA* over *SA* *more strongly* than failing to so observe

8 The screening-off assumption SO is a simplification. To see why it isn't exactly right, let's consider a version of the separate ancestry hypothesis that *guarantees* that the lineages leading to *X* and to *Y* will stray -- just a bit and only for a little while -- into the "intermediate zone" between α and β depicted in Figure 4. Given this, the *CA* and the *SA* hypothesis both entail that there are intermediate organisms. However, the two hypotheses disagree on *how many* intermediates there were. If the number of organisms in a lineage is proportional to the lineage's duration, the *CA* hypothesis says that there were lots more intermediate organisms than the *SA* hypothesis (formulated as just described) says there were. If so, when you observe an intermediate, this favors *CA* over *SA* even though both hypotheses entail that an intermediate exists. The reason this can happen is that the probability of your observing an intermediate depends, not just on *whether* such a thing exists, but on *how many* of them there are, and the two hypotheses disagree about how many. I won't attempt to replace SO with something more realistic, since the lessons I want to extract from this model would not be affected by doing so.

favors *SA* over *CA*. To evaluate this suggestion, we need a measure of strength of evidence. According to the ratio measure recommended by the Law of Likelihood, observing an intermediate favors *CA* over *SA* more strongly than failing to so observe favors *SA* over *CA* precisely when:

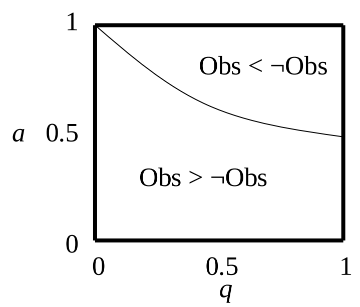
$$\frac{1}{q} > \frac{1-qa}{1-a}.$$

This simplifies to

$$\frac{1}{1+q} > a.$$

According to this criterion, each possible value of *q* puts a constraint on how large *a* is allowed to be, as shown in Figure 7. If *q* is small, practically any value for *a* will satisfy the inequality. If $a<½$, the inequality is true no matter what value *q* has. And if *q* and *a* are *both* small, the first likelihood ratio will *greatly* exceed the second. If there is a small probability of our having observed an intermediate, given that one exists, and if intermediates have a small probability of existing when the separate ancestry hypothesis is true, then observing an intermediate favors *CA* over *SA* far more profoundly than failing to so observe favors *SA* over *CA*.

Figure 7. Observing an intermediate favors *CA* over *SA*, and failing to so observe favors *SA* over *CA*, if $a>0$ and $q<1$. "Obs > ¬Obs" means that observing an intermediate favors *CA* more strongly than failing to so observe favors *SA*. Whether this is true depends on the values of the parameters *a* and *q*.



What does this show about the motto -- that *absence of evidence isn't evidence of absence*? What is true in the present context is that if you refuse to find out whether there are intermediates (and so a=0), you will certainly fail to find an intermediate, and this will be true whether the *CA* or the *SA* hypothesis is correct. Failing to find an intermediate in *this* circumstance provides zero evidence concerning the competition between *CA* and *SA*, so here the motto is correct -- absence of evidence for a common ancestor isn't evidence that there was no such thing. But this special case aside, the motto embodies an exaggeration. Suppose you *look* for intermediates and fail to find them. This outcome isn't *equally* probable under the two hypotheses if $a>0$ and $q< 1$. Entries in each column must sum to unity in Figure 6 just as they must in Figure 5. When the two parameters fall in this rather inclusive value range, failing to observe an intermediate *is* evidence against the *CA* hypothesis, contrary to the motto. What is true, without exaggeration, is that for many values of the parameters, not observing an intermediate provides *negligible* evidence favoring *SA*, compared with the much stronger evidence that observing an intermediate provides in favor of *CA*.

The fact that not observing an intermediate provides *some* evidence favoring *SA* over *CA* should not lead to creationists dancing in the streets. Testing whether two extant species have a common ancestor needs to take account of *all* the relevant evidence available. As mentioned earlier, the similarities and differences -- both phenotypic and genetic -- that characterize the two species are relevant to this enterprise. What I have described here is a piece of the total evidence – observing an intermediate fossil or failing to so observe. Even if you search for an intermediate and fail to find one, the rest of the evidence may still strongly favor *CA* over *SA*.

One part of the analysis I have offered is entirely unsurprising if the Law of Likelihood is correct. Failing to observe an intermediate *must* be evidence favoring *SA* over *CA* if observing an intermediate is evidence favoring *CA* over *SA*. This is because

$$Pr(Obs \mid CA) > Pr(Obs \mid SA) \text{ if and only if } Pr(\neg Obs \mid CA) < Pr(\neg Obs \mid SA).$$

This biconditional is a consequence of the axioms of probability; it takes on epistemic significance if the first part of the Law of Likelihood is true. Notice that this biconditional says nothing about the values of the following two likelihood ratios

$$\frac{Pr(Obs \mid CA)}{Pr(Obs \mid SA)} \qquad \frac{Pr(\neg Obs \mid SA)}{Pr(\neg Obs \mid CA)}$$

except that each must be greater than unity if the other is. I have argued that there are many cases in which the first ratio is much larger than the second. This is not a consequence of the Law of Likelihood, but it gains its epistemic significance from the second clause of that law.[9]

It may come as something of a surprise that part of what permits observing an intermediate to favor *CA* over *SA* far more than failing to so observe would favor *SA* over *CA* is that the observation would be very improbable if the *CA* hypothesis were true. To see what is going on here, consider Figure 8, which gives two sets of hypothetical values for the relevant likelihoods. In (a), observing an intermediate favors *CA* over *SA* exactly as much as failing to so observe favors *SA* over *CA*. This is not true in (b). Notice also that the likelihood ratio of *CA* to *SA* when you observe an intermediate is the same in (a) and (b); it has a value of 9. A strong

---

[9] When we fail to observe that *E* is true, we often observe that some other proposition, *E**, is true, where *E** and *E* are contraries, not contradictories. For example, suppose that when we fail to observe a fossil that is intermediate, what we do observe are fossils that are *not* intermediate. When this happens, the principle of total evidence obliges to use the logically stronger description *O(E*)* rather than ¬O(E). Although *O(E)* and ¬*O(E)* must have opposite evidential imports, *O(E)* and *O(E*)* may or may not. Even though observing fossil intermediates favors *CA* over *SA*, observing fossils that are not intermediate may or may not favor *SA* over *CA*.

asymmetry between observing and failing to observe is induced by making the observation improbable under *both* hypotheses.

Figure 8. Two examples of values for probabilities of the form *Pr(±Obs│CA)* and *Pr(±Obs│SA)*. The likelihood ratios of the hypotheses when the observation is *+Obs* are the same, but the ratios are different when *–Obs* is true.

(a)

	CA	*SA*
+Obs	0.9	0.1
-Obs	0.1	0.9

(b)

	CA	*SA*
+Obs	0.009	0.001
-Obs	0.991	0.999

In considering the evidential meaning of intermediate forms, I first considered what the *CA* and *SA* hypotheses say about the following two propositions:

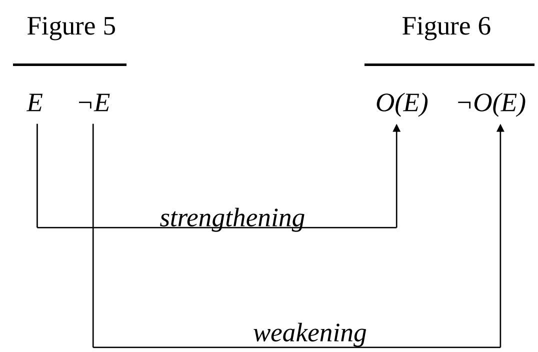
An organism intermediate between *X* and *Y* existed.
No organism intermediate between *X* and *Y* existed.

I then asked what the two hypotheses say about a second pair:

We observe that there was an organism intermediate between *X* and *Y*.
We do not observe that there was an organism intermediate between *X* and *Y*.

These two pairs of propositions, and the likelihoods they engender for the *CA* and the *SA* hypotheses, are described in Figures 5 and 6. The shift from the first pair to the second is represented in Figure 9. One part of the shift involves logically strengthening a proposition; the other involves logically weakening. The stories about toast and eggs alerted us to the possibility that logical strengthening and weakening can affect evidential meaning. The first shift, from *E* to *O(E)*, makes no difference. The second shift, from ¬*E* to ¬*O(E)*, has a quantitative effect, but no qualitative effect if $0 < a,q < 1$. By this I mean that the shift affects the magnitude of the likelihood ratio, but not whether it is greater than unity.

Figure 9. The two pairs of propositions whose evidential significance is represented in Figures 5 and 6. Moving from the first problem to the second involves strengthening one proposition, weakening the other. *E* is the proposition that there is an organism intermediate between species *X* and *Y*.

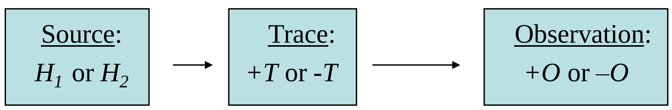


4. The General Pattern

I started this paper by saying that I wanted to investigate cases in which it is a matter of chance whether we make an observation. The probabilistic process for fossil intermediates is probabilistic twice over. There is an ontic stage, followed by an epistemic stage. The two species *X* and *Y* either have a common ancestor or they do not. Which of these alternatives is true -- a fact about their genealogies -- exerts a probabilistic influence on whether organisms ever existed whose characteristics were intermediate between those of *X* and *Y*. If intermediate organisms exist, once they die the physical processes that affect the creation and preservation of fossils take over, the result being the fossils that exist now. Then the question arises of how probable it is that a fossil now in the ground will be found and identified. The two parameters in the probability model I constructed of this two-step process pertain to different stages. Recall that q = Pr(an intermediate exists | *SA*); this describes the ontic stage. The other parameter is a = Pr(we observe an intermediate | an intermediate exists); this describes the second, epistemic, stage. For the sake of generality, let's think of this three-step process, from genealogy to intermediate organism to observed fossil, as a process that runs from Source to Trace to Observational

Result. The ontic parameter describes the relation of Source to Trace; the epistemic parameter describes the relation of Trace to Observational Result.

Figure 10. A causal chain from Source (of which H_1 or H_2 is true) to Trace (whose two states are $+T$ and $-T$) to Observation (whose two states are $+O$ and $-O$).



The model I have presented employs two assumptions that, though true in the case at hand, needn't be true in others in which we'd like to assess whether "absence of evidence isn't evidence of absence," and so the model represents a special case that needs to be generalized. The first is the assumption that one of the hypotheses considered (*CA*) *entails* that intermediate organisms exist. Although this is true in the case at hand (assuming gradualism), we want to consider the more general setting in which the competing hypotheses both confer probabilities on whether a trace exists, its being left open what values those respective probabilities have. The second assumption is that the observational result entails what state the trace was in; this is true when you observe that E is true and E is the state of the trace, but it is more general to not require this tight relationship. A more general setting of the problem is represented in Figure 10.

In this causal chain from Source to Trace to Observation, the axioms of probability entail that

$$Pr(O \mid H_1) > Pr(O \mid H_2) \text{ if and only if}$$
$$Pr(O \mid +T \,\&\, H_1)Pr(+T \mid H_1) + Pr(O \mid -T \,\&\, H_1)Pr(-T \mid H_1) >$$
$$Pr(O \mid +T \,\&\, H_2)Pr(+T \mid H_2) + Pr(O \mid -T \,\&\, H_2)Pr(-T \mid H_2).$$

If $\pm T$ screens off the H's from O, this biconditional entails that

$$Pr(O \mid H_1) > Pr(O \mid H_2) \text{ if and only if}$$
$$[Pr(O \mid +T) - Pr(O \mid -T)][Pr(+T \mid H_1) - Pr(+T \mid H_2)] > 0.$$

This implies that evidential favoring is transitive in a singly-connected causal chain with screening-off:

$$\text{If } Pr(O \mid +T) > Pr(O \mid -T) \text{ and } Pr(+T \mid H_1) > Pr(+T \mid H_2)\text{, then } Pr(O \mid H_1) > Pr(O \mid H_2).$$

If O favors $+T$ over $-T$, and $+T$ favors H_1 over H_2, then O favors H_1 over H_2. The last displayed biconditional has a second implication:

$$\text{If } Pr(O \mid T) = Pr(O \mid -T) \text{ or } Pr(+T \mid H_1) = Pr(+T \mid H_2)\text{, then } Pr(O \mid H_1) = Pr(O \mid H_2).$$

This means there are two types of "evidential breakdown" that can occur in a singly connected chain that obeys the screening-off principle. O can fail to discriminate between H_1 and H_2 for reasons pertaining to the ontic or the epistemic stages of the process. When either breakdown occurs, the two hypotheses have the same likelihood.

5. The Fine-Tuning Argument

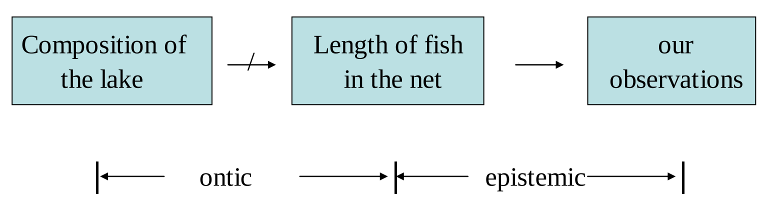
The previous discussion of the transitivity of evidence in a causal chain provides a useful setting for considering a version of the design argument for the existence of God known as the fine-tuning argument. This argument does not take the form that has become familiar from creationist attacks on evolutionary theory; it does not cite as evidence the biological fact that organisms have complex adaptive features. Rather, the argument begins with a fact from physics: the physical constants are "right," meaning that they have values that fall in the narrow range that permits life to exist. Indeed, it isn't just *life* that would have been impossible if the constants had been wrong:

> If the strong nuclear force were to have been as little as 2% stronger (relative to the other forces), all hydrogen would have been converted into helium. If it were 5% weaker, no helium at all would have formed and there would be nothing but hydrogen. If the weak

> nuclear force were a little stronger, supernovas could not occur, and heavy elements could not have formed. If it were slightly weaker, only helium might have formed. If the electromagnetic forces were stronger, all stars would be red dwarfs, and there would be no planets. If it were a little weaker, all stars would be very hot and short-lived. If the electron charge were ever so slightly different, there would be no chemistry as we know it. Carbon (^{12}C) only just managed to form in the primal nucleosynthesis (McMullin 1993, p. 378).

The suggestion is then advanced that the constants would have a higher probability of being right if our universe were produced by an intelligent designer than they'd have if the universe were produced by a mindless random process. The fine-tuning argument is a likelihood argument; the observation that the constants are right is said to favor ID over Chance.

Figure 11. Whether all the fish in the lake are longer than 10 inches, or only 50% of them are, does not affect whether the fish in the net are all over 10 inches long. The content of the net does affect what we observe.



The standard criticism of this argument invokes some version of the *anthropic principle*. The rough idea is that, since we are alive, we are bound to observe that the constants are right, regardless of whether the values of those constants were caused by ID or by Chance. We are the victims of an *observational selection effect*. Eddington (1939) provides a nice illustration of what this means. Suppose you use a net to fish in a lake and observe that all the fish in the net are over 10 inches long. At first, this observation seems to favor the hypothesis that all the fish in the lake are more than 10 inches long over the hypothesis that only 50% of them are. But then you learn that the net has holes that are 10 inches across. This makes you realize that you were bound to obtain this observation, regardless of which hypothesis about the lake is true.[10] This two-step process (Sober 2004, Bradley 2007) is depicted in Figure 11.

If you refuse to look at fossils, you'll never observe a fossil that is intermediate between species *X* and *Y*, regardless of whether *CA* or *SA* is true. If you fish with Eddington's net, you are guaranteed to observe that the net contains fish that all are over 10 inches long, regardless of whether all the fish in the lake are over 10 inches long or only 50% of them are. In the first case, you fail to make an observation while in the second, you succeed, but this difference does not matter. Both are instances of evidential breakdown. The process in which you participate guarantees that Source and Observation are related to each other by a likelihood equality. The suggestion is that the same is true of fine-tuning. If you are alive and check to see whether the physical constants fall in the narrow range that permits life to exist, you are bound to find that they do, regardless of whether the constants had their values assigned by an intelligent designer or by a chance process (Sober 2004).

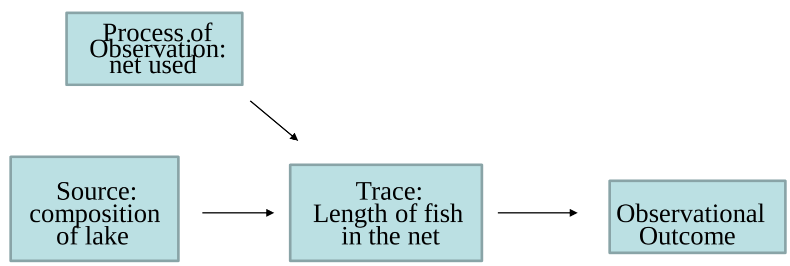
Even if this anthropic reply to the fine-tuning argument strikes you as decisive, it is important to recognize that the fine-tuning argument has intuitive appeal. It involves a causal chain from Source to Trace to Observation. The hypotheses of ID and Chance each purport to describe the Source. The Trace is characterized by whether or not the values of the physical constants are right. The observation is our observing that the values of those constants are right. The fine-tuning argument seems plausible because a transitivity argument seems correct. ID raises the probability that the constants are right,[11] and the constants' being right raises the probability that we will observe that the constants are right. Ergo, ID raises the probability of our observing that the constants are right. QED.

---

[10] I'm assuming that the net will fill with fish regardless of whether the 100% or the 50% hypothesis is true.

[11] I here set aside the objection that we don't know whether Pr(constants are right │ ID) > Pr(constants are right │ Chance). I discuss the problem of evaluating the first of these probabilities in Sober (2004) in connection with the organismic design argument. Colyvan, Garfield, and Priest (2005) have objected that the second is suspicious as well. I will assume here that the two hypotheses are formulated so as to insure that this inequality is true.

Figure 12. The contents of the net are influenced by the composition of the lake and also by the type of net used.



Those drawn to the anthropic critique of the fine-tuning argument need to explain where this transitivity argument goes wrong. I think it is defective because it fails to take account of everything that is relevant. There is more going on here than the connections just described among Source, Trace, and Observation. Before describing this extra element in the case of fine-tuning, I want to go back to Eddington’s fishing example. Figure 11 depicts it as a singly connected causal chain. Let’s modify the story a bit by imagining that you have a choice as to which type of net you’ll use. You can use a net with large (= 10 inch) holes or one with small (= 2 inch) holes. The new model is depicted in Figure 12. The contents of the net now have two causes, represented in Figure 13. Statisticians say that there is an *interaction effect* here; whether one cause raises the probability of the effect depends on the state of the other. Notice, in particular, that each cause can destroy the possibility of gaining information about the other. If you use a net with big holes, you can’t gain information about the composition of the lake by looking at the net’s contents. And if the lake contains fish that all are more than 10 inches long, you can’t tell which kind of net you used by looking at the fish that are in it. Each cause is an “epistemic switch” with respect to the other.

Figure 13. In the fishing example, the two causes of the contents of the net exhibit an *interaction effect.* Cells represent probabilities of the form Pr(the net contains fish that all are more than 10 inches long $\mid$—).

Process: Which type of net do you use?

Source: What % of fish in the lake are more than 10 inches long?

	Big Holes	Small Holes
100%	1	1
50%	1	tiny

Introducing this choice between the two nets into the fishing example makes it clear that we need to distinguish how Source and Observation are related *on average* from how they are related *in a single case*. The criterion for when the composition of the lake is on average correlated with what we see in the net is the following:

$$Pr[O(E) \mid H_{100}] > Pr[O(E) \mid H_{50}] \text{ if and only if}$$
$$Pr[O(E) \mid H_{100} \,\&\, N_L]Pr(N_L \mid H_{100}) + Pr[O(E) \mid H_{100} \,\&\, N_S]Pr(N_S \mid H_{100}) >$$
$$Pr[O(E) \mid H_{50} \,\&\, N_L]Pr(N_L \mid H_{50}) + Pr[O(E) \mid H_{50} \,\&\, N_S]Pr(N_S \mid H_{50}).$$

Here *O(E)* is the statement that you observe that the net contains fish that all are more than 10 inches long. H_{100} and H_{50} are, respectively, the hypothesis that 100% of the fish in the lake are more than 10 inches long and the hypothesis that 50% of them are. N_L is the proposition that you use a net that has large holes and N_S is the proposition that you use a net with small holes. On the assumption that the type of net you use is independent of the composition of the lake, the above biconditional simplifies to

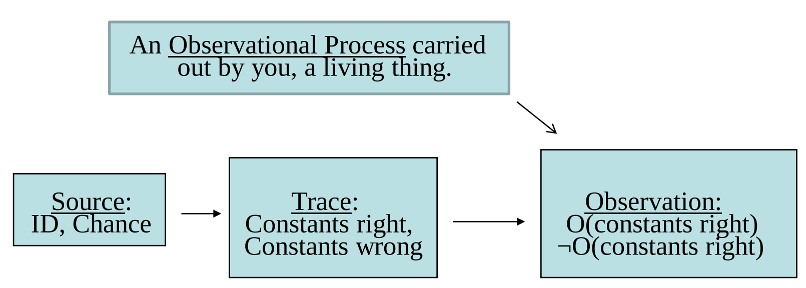
$$Pr[O(E) \mid H_{100}] > Pr[O(E) \mid H_{50}] \text{ if and only if}$$
$$Pr[O(E) \mid H_{100} \,\&\, N_L]Pr(N_L) + Pr[O(E) \mid H_{100} \,\&\, N_S]Pr(N_S) >$$
$$Pr[O(E) \mid H_{50} \,\&\, N_L]Pr(N_L) + Pr[O(E) \mid H_{50} \,\&\, N_S]Pr(N_S),$$

which further simplifies to

$$Pr[O(E) \mid H_{100}] > P[O(E) \mid H_{50}] \text{ if and only if}$$
$$Pr(N_L)\{Pr[O(E) \mid H_{100} \,\&\, N_L] - Pr[O(E) \mid H_{50} \,\&\, N_L]\} +$$
$$Pr(N_S)\{Pr[O(E) \mid H_{100} \,\&\, N_S] - Pr[O(E) \mid H_{50} \,\&\, N_S]\} > 0.$$

The first addend on the right-hand side is equal to zero, but the second one is positive (so long as $Pr(N_S) > 0$). So the observation *does* favor H_{100} over H_{50} *on average*, if which type of net you use is a matter of chance. A transitivity argument is correct in this instance. However, that argument is irrelevant if you use a net with large holes; in *that* circumstance, Source and Observation are uncorrelated. This difference between what is true on average and what is true in a specific case is an instance of Simpson's (1951) paradox.

Figure 14. In the fine-tuning argument, the state of the Source influences the state of the Trace. Your existing before you observe whether the constants are right makes a difference in whether you observe that the constants are right; the state of the Source does not.



The fine-tuning argument, like Eddington's net, involves more than a chain from Source to Trace to Observation; there is also an Observational Process that leads to the observation you make. In the case of the net, you use a net with big holes. In the fine-tuning argument, it is you, a living thing, who exists prior to the observation you make that the physical constants are right. As shown in Figure 14, your existing doesn't affect *the values* that the constants have, but it does affect whether *you'll observe* that the constants are right. The important point is that, in both examples, the process of observation prevents your observations from providing any information about the state of the Source. The relevant relationships for the fine-tuning argument are depicted in Figure 15; now there is no interaction effect.

Figure 15. In the fine-tuning argument, the impact of Source and Observational Process on the Observational Outcome involves no interaction effect. Cell entries represent Pr(you observe that the constants are right | —); $x > y$.

		Process: Do you exist before you observe whether the constants are right? Yes	No
Source:	ID	x	y
	Chance	x	y

Although I have emphasized parallelisms between fishing and fine-tuning, I recognize that there are differences. I have already noted one of them; it concerns whether there is an interaction (Figures 13 and 15). This difference can be erased by reformulating the story about fishing – suppose your choice is to use a net with big holes or to not sample from the lake at all. There seems to be a second difference. In fishing, the Process of Observation (I assume) is *independent* of the state of the Source. That is,

Pr(you use a net with big holes | 100% of the fish in the lake are more than 10 inches long) =
Pr(you use a net with big holes | 50% of the fish in the lake are more than 10 inches long).

However, it might be suggested that the Observational Process in which you are embedded and the Source are *not* independent of each other in the case of fine-tuning:

Pr(you exist | ID) > Pr(you exist | Chance).

This inequality seems to hold because ID raises the probability that there is life[12] and the existence of life raises the probability that you exist. I'll discuss this likelihood inequality later, but for now, let's suppose that it is correct. If it is, it constitutes a difference between fishing and fine-tuning; however, this difference is incidental to the epistemology. The parallelism can be made more exact by supposing that your probability of using a net with big holes is influenced by the composition of the lake. The fact remains that fishing and fine-tuning both involve an observation selection effect.

None of this is to deny that there is an inequality in the fine-tuning argument that characterizes the relationship of Source and Trace:

(I) Pr(constants are right | ID) > Pr(constants are right | Chance).

Although this inequality is true, it isn't the one on which to focus in assessing the evidence at hand. Rather, what matters is the truth of an equality. We will refine this equality shortly, but for now let us consider the following:

(Eq) Pr(you observe that the constants are right | ID & you exist) =
Pr(you observe that the constants are right | Chance & you exist).

Shifting from (I) to (Eq) involves two changes in how the problem is conceptualized. First, you are considering the probability of your *observing* that the constants are right rather than the probability that the constants are right. This change is mandated by the principle of total evidence. You are obliged to consider the logically stronger description of your evidence if this affects your assessment of the competing hypotheses. The second shift involves conditionalizing on the fact that you exist. This implements the anthropic principle -- "what we can expect to observe must be restricted by the conditions necessary for our presence as observers" (Carter 1974). Whereas the principle of total evidence is clear enough for present purposes, the anthropic principle needs to be examined more carefully.

[12]As mentioned in footnote 11, I grant this for the sake of argument.

6. The Anthropic Principle

In the case of Eddington's net, it is not enough to conditionalize on the proposition that "we are observers" or on the proposition that "we will observe what the net contains." We also must take account of the process by which we came to make our observations – in particular, the net we actually used. Within a likelihood framework, the following is a natural formulation of the principle we seek:

> *The Anthropic Principle*: An agent applying the Law of Likelihood to determine how the observational result E bears on the hypotheses H_1 and H_2 should compare $Pr[O(E) \mid H_1 \& P]$ and $Pr[O(E) \mid H_2 \& P]$, where P provides as complete a description as the agent possesses of the process by which he or she came to observe that E is true.

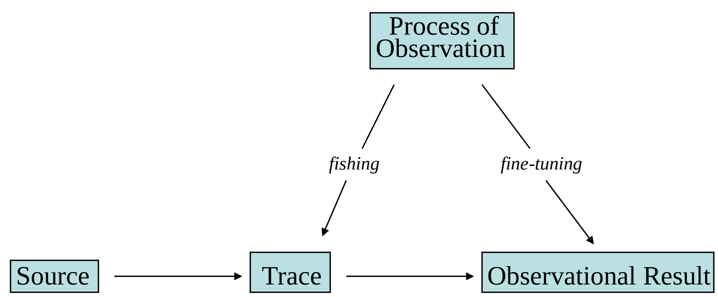
When the process of observation is stochastic, the anthropic principle entails that you need to be as specific as you can; for example, if you know that it is a matter of chance whether you use a net with big holes or a net with small ones, and also that you end up using a net with big holes, you should conditionalize not just on the chancy fact, but on the type of net you actually happen to use. Not just any old description of "the process of observation" will do.

Although the anthropic principle enjoins agents to be maximally specific, "the process of observation" cannot be taken to include *all* factors that influence what the observational result is. For example, the hypotheses under test (H_1 versus H_2) may make a difference in the probability of the observations. But, if H_1 is true, you can't include that proposition in P, since doing so will have the result that one of the conditional probabilities you want to consider is not defined (assuming that H_1 and H_2 are incompatible).[13] Although we need to distinguish the process of observation from the hypotheses under test, this leaves open, as noted before, that the two may be correlated. What is required is that P be *logically* independent of H_1 and of H_2.

---

[13] Even when you *know* which of the hypotheses is true, it still is possible that a given bit of evidence discriminates between them, and not always in favor of the true hypothesis.

Figure 16. The general format that needs to be used to assess whether there is an observation selection effect. In addition to the process from Source to Trace to Observational Result, there is also the Process of Observation. This can impinge on the Trace, the Observational Result, or on both. Fishing with Eddington's net and the fine-tuning argument are located as shown.



What about the observational outcome *O(E)*? Can this (and propositions that entail it) be inserted into the description of the process of observation (*P*)? Doing so engenders a likelihood equality:

$$Pr[O(E) \mid H_1 \& O(E)] = Pr[O(E) \mid H_2 \& O(E)] = 1.$$

To *always* insert *O(E)* into *P* is to lapse into evidential nihilism — to embrace the view that observations never discriminate between competing hypotheses. I am no nihilist. Though observation selection effects *sometimes* occur, they are not ubiquitous. What, then, settles whether propositions entailing *O(E)* get inserted into the description of the process *P*? We are working with a relationship among four elements – the hypotheses under test, the state of a trace, the process of observation, and the observational outcome, as shown in Figure 16. The difference between the process leading to an observation and the observation itself is often clear enough; the problem is that the distinction between process and product becomes blurred when there is an observation selection effect.

This problem resembles one that has been discussed in connection with the question of how the thesis of determinism ought to be described (Berofsky 1971, p. 161; Earman 1986, p. 14). The rough idea of (forwards) determinism is that any two systems that are in the same (total) state at one time must be in the same state forever after. But what counts as a proper description of a system's state? Suppose it is true at t_1 that the system will be in a given state at t_2. The difficulty is that determinism will be true *by definition* if this fact is counted as describing the state of the system at t_1. Although it is probably impossible to give a formal criterion that defines what constitutes the state of a system at a time, this does not show that the question of whether determinism is true is a pseudo-problem. The same goes for the question of whether a given observation is subject to an observation selection effect.

I have two suggestions for organizing our thinking about observation selection effects. The first is to impose different time indices on the events we wish to describe. In the fishing example, the composition of the lake is set at t_1 and remains unchanged thereafter. At t_2, I choose a net and place it in the water. At t_3, I draw up the net and observe that all the fish in it are more than 10 inches long. In fine-tuning, ID or Chance set the values of the physical constants at t_1 and those values remain unchanged thereafter. At t_2, I come into existence. At t_3, I observe that the constants are right. The general idea is that we can avoid conflating Source, Observational Process, and Observational Outcome by relegating each to its own temporal period. Waiving Aristotelian worries about the truth values of future contingents, I suppose it is true at t_1 and t_2 that I observe at t_3 that the constants are right. However, that fact does not count as a proper description of the state of the system at those earlier times.

The second suggestion is that we take seriously the fact that our surviving from moment to moment is a matter of probability. We think of ourselves as deciding today what we will do tomorrow. This is not a mistake, but what we actually achieve is something modest; we make it true today that we will do various things tomorrow if we survive and if other favorable conditions conspire (e.g., we don't forget our earlier resolve). "Deciding today to do *X* tomorrow" is not factive; it does not entail that we will do *X* tomorrow.

The likelihood equality (Eq) can now be sharpened by being more explicit about the timing of events:

Pr(I observe at t_3 that the constants are right | the values of the constants are set by an intelligent designer at t_1 & I am alive at t_2) =
Pr(I observe at t_3 that the constants are right | the values of the constants are set by a chance process at t_1 & I am alive at t_2).

Weisberg (2005) defends a different representation of the fine-tuning argument. He thinks that the anthropic principle, properly understood, tells us to conditionalize on a different set of background assumptions. According to Weisberg, there is no observation selection effect and the correct representation of the argument is a likelihood inequality:

> Pr(I observe at t_3 that the constants are right | the values of the constants are set by an intelligent designer at t_1 & If I am alive at t_3, I will observe whether the constants are right) >
> Pr(I observe at t_3 that the constants are right | the values of the constants are set by a chance process at t_1 & If I am alive at t_3, I will observe whether the constants are right).

Which of these indented statements is the relevant one to consider – Weisberg's inequality or my equality? Notice that Weisberg does not conditionalize on the fact that I am alive at t_2. I do so because I think this simple fact is part of the process leading to my observation at t_3. It is just as much part of the process as the choice of a net with which to fish. Given that I am alive at t_2, the constants must be right at t_2. And if they are right at t_2, they also will be right at t_3. The constants, I take it, do not change values from one moment to the next.[14] There is an observation selection effect. It isn't that my deciding at t_2 to observe at t_3 whether the constants are right insures that I'll observe at t_3 that they are; I may die in the interval or my plans may change. Rather, the point is that my being alive at t_2 insures that the constants are right at t_3. Given this, my probability of observing at t_3 that the constants are right is the same, regardless of whether ID or Chance is true.

Suppose a baby is born at t_3 and immediately observes that the constants are right (some smart baby, you may be thinking). This baby does not believe at t_3 that she was alive at t_2, so she can hardly be expected to conditionalize on that fact. Does this rescue the fine-tuning argument? No. The baby existed prenatally and her parents, grandparents, and other ancestors were alive

[14] I take it that the fine-tuning argument assumes that the physical constants do not change values during the duration of our universe; the question is what gave those constants their unchanging values. The argument can of course be reformulated so that the constants are liable to change. If the Chance* hypothesis entails that the constants can flip-flop between right and wrong from moment to moment, it is arguable that ID is more likely than Chance*. However, the fact remains that ID is not more likely than Chance.

before that. She trails that history behind her like a net. She is the victim of an observation selection effect, even though she does not know it.[15]

7. Observation Selection Effects

The anthropic principle is a pragmatic principle, giving advice to agents about which probabilities they should use to interpret their evidence. The principle allows that you and I might differ in the background knowledge we possess and therefore interpret the same evidence differently. If you know that a net with large holes was used, you will conclude that the contents of the net provide no evidence about the contents of the lake. If I know only that it was a matter of chance which type of net was used, I will conclude that the contents of the net do provide evidence. You and I are doing the best we can with the different information we possess.

There is an objective concept that needs to be defined here, one that isn't relative to an agent's background beliefs. If you use a net with big holes, you are the victim of an observation selection effect, regardless of whether you know that this is the case. What matters is the process by which you obtained your observations, not what you happen to know or believe about it. A natural way to capture this concept is by envisioning a complete description of the process of observation:

> An agent who uses the observation that E is true to evaluate the competition between H_1 and H_2 is the victim of an observation selection effect if and only if $\frac{\Pr[O(E) / H1 \& B]}{\Pr[O(E) / H2 \& B]} \neq \frac{Pr[O(E) \mid H_1 \& P]}{Pr[O(E) \mid H_2 \& P]}$, where B is the set of background beliefs the agent uses and P provides a complete description of the process by which the agent obtained their observation.[16]

When your choice of net is a matter of chance, though you happen to use a net with big holes, the inequality is true because the first ratio is greater than one, while the second is equal to unity. The same is true of fine-tuning. Observation selection effects are instances of *sampling bias*.[17]

---

[15] Cases of amnesia can be used to make the same point. Is this a connection with the sleeping beauty problem?

[16] Note the reliance of the likelihood ratio in this definition. Other measures of strength of evidence would require other definitions.

[17] I distinguish qualitative and quantitative observation selection effects in Sober (2004). Fishing and fine-tuning are instances of the former, not just the latter.

8. The Firing Squad

The fine-tuning argument is a rabbit/duck. Put the argument next to the story about Eddington's net and the argument seems to be vitiated by an observation selection effect. But place the argument next to Leslie's (1989) story about a firing squad, and the argument seems to be correct; now no observation selection effect seems to be in play. The gestalt switch is striking.

Leslie describes a prisoner who stands before a firing squad. The squad fires. To the prisoner's surprise, he finds that he is still alive. He then uses the fact that he is alive to evaluate the following two hypotheses. The ID hypothesis says that the members of the firing squad intended to spare him. The Chance hypothesis says that the members of the squad fired their weapons in randomly selected directions. Leslie thinks it is clear that the prisoner's being alive is evidence for ID over Chance. For those who agree, the relevant point is the following likelihood inequality (stated in the 1st person, from the prisoner's point of view):

$$\Pr(\text{I am alive at } t_3 \mid \text{ID at } t_1) > \Pr(\text{I am alive at } t_3 \mid \text{Chance at } t_1).$$

Leslie maintains that the prisoner would be making a mistake if he set aside this inequality and focused instead on the following equality:

$$\Pr(\text{I am alive at } t_3 \mid \text{ID at } t_1 \text{ \& I observe at } t_3 \text{ whether I am alive})$$
$$= \Pr(\text{I am alive at } t_3 \mid \text{Chance at } t_1 \text{ \& I observe at } t_3 \text{ whether I am alive}).$$

Leslie concludes that since the prisoner reasons correctly by invoking the likelihood inequality, the fine-tuning argument is correct and criticisms of it that appeal to the anthropic principle are misguided.

I argued in my 2004 paper that the prisoner is in the grip of an observational selection effect and so, when the relevant likelihoods are used to evaluate his evidence, a likelihood equality arises and he must conclude that his survival does not discriminate between ID and Chance.[18] Weisberg (2005) and several other patient friends have persuaded me that I was

[18] I also argued in that paper that the prisoner has other sources of information that permit him to interpret his survival in a way that bypasses questions about likelihoods; this complication can be ignored here.

mistaken. But I still think there is an observation selection effect in the fine-tuning argument. If the prisoner is right, how can the fine-tuner be wrong? Careful attention to time indices draws our attention to the following:

Pr(I observe at t_3 that I am alive | the firing squad decides at t_1 that it will spare my life when it fires at t_2 & I am alive at t_2 when the squad fires) >
Pr(I observe at t_3 that I am alive | the firing squad decides at t_1 that it will fire in randomly chosen directions at t_2 & I am alive at t_2 when the squad fires).

There is a difference between the firing squad and fine-tuning. I missed it in Sober (2004), thinking that there is an observation selection effect in both. Weisberg missed it too, thinking that there is an observation selection effect in neither. The prisoner's being alive at t_2 does not screen off ID and Chance from his observing at t_3 that he is alive. But my being alive at t_2 does screen off ID and Chance from my observing at t_3 that the constants are right.

9. Another Wrinkle

Friends of fine-tuning need to provide a principled basis for denying that the process of observation (*P*), properly understood, screens off ID and Chance from the observation that the physical constants are right; it is not enough to point to supposed analogies with other examples. Here I want to comment on a second line of argument that they sometimes advance. Even if observing that the constants are right does not favor ID over Chance, maybe there is another observation in the neighborhood that does. What about the fact, briefly mentioned in Section 5, that you exist at t_2? Even if this is background knowledge relative to your observing at t_3 that the constants are right, isn't the fact that you exist at t_2 itself evidence that favors ID over Chance? Or perhaps you should consider all the observations you have ever made, starting with your first perceptions (this is the case of the newborn baby who grows up). Isn't that totality more probable under the ID hypothesis than it is under the hypothesis of Chance?

Here we have left the fine-tuning argument behind and have shifted to another sort of alleged evidence. Let us consider it briefly. I have already mentioned that the newborn baby can be the victim of an observation selection effect even if her reasoning conforms to the dictates of the anthropic principle. But there is a second issue that needs to be faced. We need to consider whether your existing really does have a higher probability under the ID hypothesis than it does under the hypothesis of Chance. Maybe an intelligent designer, if such a being existed, would have insured that you would not exist. If so, you'd have a higher probability of existing under the Chance hypothesis. The problem is even more striking when we consider the totality of observations you have made, starting from day one. Maybe an intelligent designer, if such a being existed, would have insured that at least one of these observed events would not have occurred. If so, your total body of observations would have a probability of zero under the ID hypothesis, but would be nonzero under the hypothesis of Chance. I do not claim that these probabilities under the ID hypothesis really are zero, but that there is a serious question about why we are entitled to assume that they are not.[19] In any event, my subject here is fine-tuning. I do not claim to have provided a recipe for debunking all possible versions of the design argument.

10. Concluding Comments

It is a consequence of the axioms of probability theory that

$$Pr[O(E) \mid H_1 \& P] > Pr[O(E) \mid H_2 \& P] \text{ if and only if}$$
$$Pr[notO(E) \mid H_1 \& P] < Pr[(notO(E) \mid H_2 \& P].$$

Here H_1 and H_2 are the hypotheses under test, *O(E)* is the proposition that you observed that *E* is true, and *P* describes the process leading up to that observational outcome. It is not a mathematical theorem that the inequalities on the left and right of this biconditional say anything of interest epistemologically, but, according to the Law of Likelihood, they do. I have not discussed why the qualitative part of the Law of Likelihood should be accepted; rather, this paper is predicated on the assumption that it should be. If the Law is wrong, both my evaluation of the slogan "absence of evidence isn't evidence of absence" and my formulation of the anthropic principle need to be rethought. The same point holds, of course, for my use of the quantitative part of the Law of Likelihood.

The biconditional displayed above, when coupled with the Law of Likelihood, entails that there is a symmetry between observing and failing to observe. Evolutionists often maintain that observing a fossil intermediate is evidence *for* common ancestry but that failing to so observe isn't evidence *against*. Creationists are the mirror image; they often maintain that failing to observe a fossil intermediate is evidence *against* common ancestry but that finding such fossils isn't evidence *for*. Both parties are wrong. If the circumstances of observation render *O(E)*

[19] This is an important question to pose for the organismic design argument, as mentioned in footnote 11.

evidentially informative, those circumstances also render *notO(E)* informative. Although this biconditional expresses a *qualitative* symmetry, it does not entail that there also is a *quantitative* symmetry. When absence of evidence is evidence of absence, it does not follow that absence of evidence is *strong* evidence of absence. It is perfectly possible that *O(E)* provides strong evidence favoring H_1 over H_2 while *notO(E)* provides only weak evidence favoring H_2 over H_1. I explored this possibility in connection with the issue of fossil intermediates by comparing two likelihood ratios.

One theme that connects the discussions of absence of evidence and of observation selection effects is the importance of attending to shifts between logically stronger and logically weaker descriptions of the evidence. In the case of fossils, it makes all the difference whether you consider "there is no fossil intermediate between species *X* and *Y*" or "no fossil that is intermediate between *X* and *Y* has been observed." Here we move from a logically stronger to a logical weaker evidence statement. In the fine-tuning argument, we shifted from "the constants are right" to "we observe that the constants are right;" here the shift is from logically weaker to logically stronger. The principle of total evidence, a mainstay of Bayesianism, says that logically stronger descriptions of the evidence must be used if they make a difference in the interpretation of evidence. The formulation offered here of the anthropic principle makes that principle a close cousin of the principle of total evidence. Both are pragmatic principles that give advice concerning which probabilities are epistemically relevant. Their difference consists in the fact that one of them focuses on what goes to the left of the conditional probability sign while the other concerns what goes to the right in expressions of the form $Pr[O(E) \mid H\&P]$.

A further point of contact between fossils on the one hand and fishing, fine-tuning, and firing squads on the other concerns the correctness of transitivity inferences about evidence. In a causal chain from Source to Trace to Observation in which there is screening off, if the Observation provides evidence about the Trace and the Trace provides evidence about the Source, then the Observation provides evidence about the Source. This transitivity principle indicates that there can be two sorts of evidential breakdown. If you refuse to look at fossils, you'll certainly fail to observe a fossil intermediate regardless of whether the common ancestry or the separate ancestry hypothesis is true. And if you fish with a net that has big holes, you'll observe that the fish in your net are all more than 10 inches long regardless of the composition of the lake. In the first case you fail to observe something while in the second you do make an observation, but in both cases the upshot fails to discriminate between the competing hypotheses.

Evidential transitivity seems plausible and it is part of the reason that the fine-tuning argument seems right. But this is a mistake. When there are *two* nets one might use in the fishing example (Figure 13), it is important to distinguish whether Observation and Source are correlated *on average* from whether they are correlated in the specific situation in which one has made one's observations. In similar fashion, ID raises the probability that the constants are right (let us suppose) and the constants' being right raises the probability that we will observe that they are. This is true but misses the point, or so the anthropic principle asserts. We need to focus on what occurs during the observational process, just as is true in the case of fishing.

If the motto – that absence of evidence isn't evidence of absence -- is often wrong, why does it persist? One reason is that there is a special case in which it is exactly right and this case is especially vivid – if you refuse to look at fossils, you'll fail to observe a fossil intermediate regardless of whether the common or the separate ancestry hypothesis is true. The other reason the motto persists is that when it is false, it is often close to being true – it involves an exaggeration that is slight. When the parameters q and a both have small values, observing a fossil intermediate *strongly* favors the common ancestry hypothesis over the hypothesis of separate ancestry, whereas failing to so observe *very weakly* favors separate ancestry over common ancestry. When something makes a small difference, it often seems harmless to say that it makes no difference.[20] If you ask a biologist whether the torpedo shape exhibited by sharks and dolphins is evidence that they have a common ancestor, you may receive the answer that it does not. The biologist may explain that you'd expect large aquatic predators to evolve this shape whether or not they have a common ancestor; the shape is highly adaptive, and that is why it provides no evidence for common ancestry. In the *Origin*, Darwin (1859, p. 424) describes this line of reasoning, but he avoids overstatement:

> … adaptive characters, although of the utmost importance to the welfare of the being, are almost valueless to the systematist. For animals belonging to two most distinct lines of descent, may readily become adapted to similar conditions, and thus assume a close external resemblance; but such resemblances will not reveal – will rather tend to conceal their blood-relationship to their proper lines of descent.

---

[20] This is what many Bayesians say about the raven paradox: observing a black raven and observing a non-black non-raven both confirm "All ravens are black," but the former provides strong confirmation while the latter provides weak. Saying that non-black non-ravens don't confirm at all is therefore a mild exaggeration. See Eells (1982, p. 61) for discussion and references.

Darwin says that adaptive similarities are *almost* valueless, not that they are *completely* so. Darwin's wording isn't awkward, though maybe his prose would have been more striking if he had dropped "almost" and used "completely" instead. Be that as it may, the slogan "absence of evidence often provides only weak evidence of absence" does not fall trippingly off the tongue. Imagine writing those words on a T-shirt that you want to sell. Perhaps there is no pithy aphorism that exactly captures what is true about absence of evidence. We therefore find it natural to opt for brevity over exactness when we say that absence of evidence isn't evidence of absence.

Acknowledgements

My thanks to Matthew Barker, Bernard Berofsky, Darren Bradley, Juan Comesaña, James Crow, Daniel Hausman, Steven Leeds, Gregory Mougin, Robert Northcott, Gregory Novack, Carolina Sartorio, Yehuda Schnall, Christopher Stephens, Michael Strevens, Martin Thomson-Jones, Michael Titelbaum, Bas Van Fraassen, Peter Vranas, Paul Weirich, Jonathan Weisberg, and Roger White for useful discussion. I am especially grateful to Martin Barrett.

References

Berofsky, B. (1971). *Determinism.* Princeton: Princeton University Press.

Bradley, D. (2007). Bayesianism and Self-Locating Beliefs, or Tom Bayes meets John Perry. Ph.D. dissertation, Stanford University.

Carroll, S. (2005). *Endless Forms Most Beautiful – the New Science of Evo Devo.* New York: Norton.

Carter, B. (1974). Large Number Coincidences and the Anthropic Principle in Cosmology. In M.S. Longair (ed.), *Confrontation of Cosmological Theories with Observational Data* (pp. 291-298). Dordrecht: Reidel.

Colyvan, M., Garfield, J. and Priest, G. (2005). Problems with the Argument from Fine-Tuning. *Synthese*, 145, 325-338.

Darwin, C. (1859). *On the Origin of Species*. London: Murray.

Earman, J. (1986). *A Primer on Determinism.* Doredrecht: Reidel.

Earman, J. (1992). *Bayes or Bust – A Critical Examination of Bayesian Confirmation Theory.* Cambridge: MIT Press.

Eddington, A. (1939). *The Philosophy of Physical Science*. Cambridge: Cambridge University

Press.

Edwards, A. (1972). *Likelihood*. Cambridge: Cambridge University Press.

Eells, E. (1982). *Rational Decision and Causality*. Cambridge: Cambridge University Press.

Fitelson, B. (1999). The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity. *Philosophy of Science*, 66, S362-S378.

Hacking, I. (1965). *The Logic of Statistical Inference*. Cambridge: Cambridge University Press.

Leslie, J. (1989). *Universes*. London: Routledge.

McMullin, E. (1993). Indifference Principle and Anthropic Principle in Cosmology. *Studies in the History and Philosophy of Science*, 24, 359-389.

Royall, R. (1997). *Statistical Evidence – a Likelihood Paradigm*. Boca Raton, FL: Chapman and Hall.

Sagan, C. (1997). *The Demon Haunted World – Science as a Candle in the Dark.* New York: Random House.

Simpson, E.H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society B*, 13, 238-241.

Sober, E. (2004). The Design Argument. In W. Mann (ed.), *The Blackwell Guide to Philosophy of Religion* (pp. 117-147), New York: Blackwell.

Sober, E. (2008). *Evidence and Evolution – the Logic Behind the Science.* Cambridge: Cambridge University Press.

Walton, D. (1996). *Arguments from Ignorance*. University Park, PA: Penn State Press.

Weisberg, J. (2005). Firing Squads and Fine Tuning – Sober on the Design Argument. *British Journal for the Philosophy of Science*, 56, 809-821.