

# **The Explanatory Filter:**

## **A three-part filter for understanding how to separate and identify cause from intelligent design**

**An excerpt from a paper presented at the 1996 Mere Creation conference, originally titled "Redesigning Science."**

---

**William A. Dembski, Ph.D.**

Center for the Philosophy of Religion, University of Notre Dame

What is science going to look like once Intelligent Design succeeds? To answer this question we need to be clear what we mean by Intelligent Design. Intelligent Design is not repackaged creationism, nor religion masquerading as science. Intelligent Design holds that intelligent causation is an irreducible feature of the bio-physical universe, and furthermore that intelligent causation is empirically detectable. It is unexceptionable that intelligent causes can do things which unintelligent causes cannot. Intelligent Design provides a method for distinguishing between intelligent and unintelligent causes, and then applies this method to the special sciences.

Hardly a dubious innovation, Intelligent Design formalizes and makes precise something we do all the time. All of us are all the time engaged in a form of rational activity which, without being tendentious, can be described as inferring design. Inferring design is a perfectly common and well-accepted human activity. People find it important to identify events that are caused through the purposeful, premeditated action of an intelligent agent, and to distinguish such events from events due to either law or chance. Intelligent Design unpacks the logic of this everyday activity, and applies it to questions in science. There's no magic, no vitalism, no appeal to occult forces here. Inferring design is widespread, rational, and objectifiable. The purpose of this paper is to formulate Intelligent Design as a scientific theory.

The key step in formulating Intelligent Design as a scientific theory is to delineate a method for detecting design. Such a method exists, and in fact, we use it implicitly all the time. The method takes the form of a three-stage Explanatory Filter. Given something we think might be designed, we refer it to the filter. If it successfully passes all three stages of the filter, then we are warranted asserting it is designed. Roughly speaking the filter asks three questions and in the following order: (1) Does a law explain it? (2) Does chance explain it? (3) Does design explain it?

To see how the filter works in practice, consider the case of Nicholas Caputo. Back in 1985 Nicholas Caputo was brought before the New Jersey Supreme Court. The Republican party had filed suit against him, claiming Caputo had consistently rigged the ballot line in Essex County, New Jersey where he was county clerk. It is a known fact that first position on a ballot increases one's chances of winning an election. Since in every instance but one Caputo positioned the Democrats first on the ballot line, the

Republicans argued that in selecting the order of ballots Caputo had intentionally favored his own Democratic party. In short, the Republicans claimed Caputo had cheated.

The question then before the New Jersey Supreme Court was, Did Caputo actually rig the order, or was it without malice and forethought on his part that the Democrats happened 40 out of 41 times to appear first on the ballot? Since Caputo denied wrongdoing, and since he conducted the drawing of ballots so that witnesses were unable to observe how he actually did draw the ballots, determining whether Caputo did in fact rig the order of ballots becomes a matter of evaluating the circumstantial evidence connected with this case. How then is this evidence to be evaluated?

In determining how to explain the remarkable coincidence of Nicholas Caputo selecting the Democrats 40 out of 41 times to head the ballot line, the court had three options to consider:

### **Law**

Unbeknownst to Caputo, he was not employing a reliable random process to determine ballot order. Caputo was in the position of someone who thinks she is flipping a fair coin when in fact she is flipping a double-headed coin. Just as flipping a double-headed coin is going to yield a long string of heads, so Caputo, using his faulty method for ballot selection, generated a long string of Democrats coming out on top.

### **Chance**

In selecting the order of political parties on the state ballot, Caputo employed a reliable random process that did not favor one political party over another. The fact that the Democrats came out on top 40 out of 41 times was simply a fluke. It occurred by chance.

### **Design**

Caputo, knowing full well what he was doing and intending to aid his own political party, purposely rigged the ballot line selection process so that the Democrats would consistently come out on top. In short, Caputo cheated.

The first option—that Caputo chose poorly his procedure for selecting ballot lines, so that instead of genuinely randomizing the ballot order, it just kept putting the Democrats on top—was dismissed by the court because Caputo himself had claimed to use a randomization procedure in selecting ballot lines. And since there was no reason for the court to think that Caputo's randomization procedure was at fault, the key question therefore became whether Caputo actually put this procedure into practice when he made the ballot line selections, or whether he purposely circumvented this procedure in order for the Democrats consistently to come out on top. And since Caputo's actual drawing of the capsules was obscured to witnesses, it was this question that the court had to answer.

With the law explanation eliminated, the court next decided to dispense with the chance explanation. Having noted that the chances of picking the same political party 40 out of 41 times were less than 1 in 50 billion, the court concluded that "confronted with these

odds, few persons of reason will accept the explanation of blind chance." Now this certainly seems right. Nevertheless, a bit more needs to be said. The problem is that the exceeding improbability is by itself not enough to preclude something from happening by chance.

Invariably, what is needed to eliminate chance is that the event in question conform to a *pattern*. Not just any pattern will do, however. Some patterns can legitimately be employed to eliminate chance whereas others cannot.

A bit of terminology will prove helpful here. The "good" patterns will be called *specifications*. Specifications are the non-*ad hoc* patterns that can legitimately be used to eliminate chance and warrant a design inference. In contrast, the "bad" patterns may be called *fabrications*. Fabrications are the *ad hoc* patterns that cannot legitimately be used to eliminate chance.

By selecting the Democrats to head the ballot 40 out of 41 times, Caputo appears to have participated in an event of probability less than 1 in 50 billion. Yet, exceedingly improbable things happen all the time. The crucial question therefore is whether this event is also specified—does this event follow a non-*ad hoc* pattern so that we can legitimately eliminate chance? But of course, the event is specified: that Caputo is a Democrat, that it is in Caputo's interest to see the Democrats appear first on the ballot, that Caputo controls the ballot lines, and that Caputo would by chance be expected to assign Republicans top ballot line as often as Democrats all conspire to specify Caputo's ballot line selections, and render his selections incompatible with chance. No one to whom I have shown this example draws any other conclusion than design, to wit, Caputo cheated.

In the trial of Nicholas Caputo the New Jersey Supreme Court employed the Explanatory Filter, first rejecting a law explanation, then a chance explanation, and finally inferring a design explanation.

At the first stage, the filter determines whether a law can explain the thing in question. Law thrives on replicability, yielding the same result whenever the same antecedent conditions are fulfilled. Clearly, if something can be explained by a law, it better not be attributed to design. Things explainable by a law are therefore eliminated at the first stage of the Explanatory Filter.

Suppose, however, that something we think might be designed cannot be explained by any law. We then proceed to the second stage of the filter. At this stage the filter determines whether the thing in question might not reasonably be expected to occur by chance. What we do is posit a probability distribution, and then find that our observations can reasonably be expected on the basis of that probability distribution. Accordingly, we are warranted attributing the thing in question to chance. And clearly, if something can be explained by reference to chance, it better not be attributed to design. Things explainable by chance are therefore eliminated at the second stage of the Explanatory Filter.

Suppose finally that no law is able to account for the thing in question, and that any plausible probability distribution that might account for it does not render it very likely. Indeed, suppose that any plausible probability distribution that might account for it renders it exceedingly unlikely. In this case we bypass the first two stages of the Explanatory Filter and arrive at the third and final stage. It needs to be stressed that this third and final stage does not automatically yield design—there is still some work to do. Vast improbability only purchases design if, in addition, the thing we are trying to explain is specified.

The third stage of the Explanatory Filter therefore presents us with a binary choice: attribute the thing we are trying to explain to design if it is specified; otherwise, attribute it to chance. In the first case, the thing we are trying to explain not only has small probability, but is also specified. In the other, it has small probability, but is unspecified. It is this category of specified things having small probability that reliably signals design. Unspecified things having small probability, on the other hand, are properly attributed to chance.

The Explanatory Filter faithfully represents our ordinary practice of sorting through things we alternately attribute to law, chance, or design. In particular, the filter describes

how copyright and patent offices identify theft of intellectual property

how insurance companies prevent themselves from getting ripped off

how detectives employ circumstantial evidence to incriminate a guilty party

how forensic scientists are able reliably to place individuals at the scene of a crime

how skeptics debunk the claims of parapsychologists

how scientists identify cases of data falsification

how NASA's SETI program seeks to identify the presence of extra-terrestrial life, and

how statisticians and computer scientists distinguish random from non-random strings of digits.

Entire industries would be dead in the water without the Explanatory Filter. Much is riding on it. Using the filter, our courts have sent people to the electric chair. Let us now see why the filter works.

### Why the Filter Works

The filter is a criterion for distinguishing intelligent from unintelligent causes. Here I am using the word "criterion" in its strict etymological sense as a method for deciding or

judging a question. The Explanatory Filter is a criterion for deciding when something is intelligently caused and when it isn't. Does it decide this question reliably?

As with any criterion, we need to make sure that whatever judgments the criterion renders correspond to reality. A criterion for judging the quality of wines is worthless if it judges the rot-gut consumed by winos superior to a fine French Bordeaux. The reality is that a fine French Bordeaux is superior to the wino's rot-gut, and any criterion for discriminating among wines better indicate as much.

Or consider medical tests. Any medical test is a criterion. A perfectly reliable medical test would detect the presence of a disease whenever it is indeed present, and fail to detect the disease whenever it is absent. Unfortunately, no medical test is perfectly reliable, and so the best we can do is keep the proportion of false positives and false negatives as low as possible.

All criteria, and not just medical tests, face the problem of false positives and false negatives. A criterion attempts to classify individuals with respect to a target group (in the case of medical tests, those who have a certain disease). When the criterion classifies an individual who should not be there in the target group, it commits a false positive. Alternatively, when the criterion fails to classify an individual who should be there in the target group, it commits a false negative. Take medical tests again. A medical test checks whether an individual has a certain disease. The target group comprises all those individuals who actually have the disease. When the medical test classifies an individual who doesn't have the disease with those who do, it commits a false positive. When the medical test classifies an individual who does have the disease with those who do not, it commits a false negative.

When the Explanatory Filter fails to detect design in a thing, can we be sure no intelligent cause underlies it? The answer to this question is No. For determining that something is not designed, the Explanatory Filter is not a reliable criterion. False negatives are a problem for the Explanatory Filter. This problem of false negatives, however, is endemic to detecting intelligent causes. One difficulty is that intelligent causes can mimic law and chance, thereby rendering their actions indistinguishable from these unintelligent causes. It takes an intelligent cause to know an intelligent cause, but if we don't know enough, we'll miss it.

Intelligent causes can do things that unintelligent causes cannot, and can make their actions evident. When for whatever reason an intelligent cause fails to make its actions evident, we may miss it. But when an intelligent cause succeeds in making its actions evident, we take notice. This is why false negatives do not invalidate the Explanatory Filter. The Explanatory Filter is fully capable of detecting intelligent causes intent on making their presence evident.

And this brings us to the problem of false positives. Even though the Explanatory Filter is not a reliable criterion for *eliminating* design, it is, I argue, a reliable criterion for *detecting* design. The Explanatory Filter is a net. Things that are designed will

occasionally slip past the net. We would prefer that the net catch more than it does, omitting nothing due to design. But given the ability of design to mimic unintelligent causes and the possibility of our own ignorance passing over things that are designed, this problem cannot be fixed. Nevertheless, we want to be very sure that whatever the net does catch includes only what we intend it to catch, to wit, things that are designed.

I argue that the explanatory filter is a reliable criterion for detecting design. Alternatively, I argue that the Explanatory Filter successfully avoids false positives. Thus whenever the Explanatory Filter attributes design, it does so correctly.

Let us now see why this is the case. I offer two arguments. The first is a straightforward inductive argument: in every instance where the Explanatory Filter attributes design, and where the underlying causal story is known, it turns out design actually is present; therefore, design actually is present whenever the Explanatory Filter attributes design.

My second argument for showing that the Explanatory Filter is a reliable criterion for detecting design may now be summarized as follows: the Explanatory Filter is a reliable criterion for detecting design because it coincides with how we recognize intelligent causation generally. In general, to recognize intelligent causation we must observe a choice among competing possibilities, note which possibilities were not chosen, and then be able to specify the possibility that was chosen.

### **The Relevance to Biology**

One thing is clear. Creationists and evolutionists alike feel the force of design. At some level they are all responding to it. This is true even of those who, unlike Dawkins, think that life is extremely unlikely to occur by chance in the known physical universe, but who nevertheless agree with Dawkins that life is properly explained without reference to design. Here I have in mind advocates of the Anthropic Principle, like Barrow and Tipler (1986), who posit an ensemble of universes so that life, though highly improbable in our own little universe, is nevertheless virtually certain to have arisen at least once in the many- many universes that constitute the ensemble of which our universe is a member. This move allows Barrow and Tipler to vastly multiply their probabilistic resources, and thus vastly lower their probability for the origin of life on earth.

There remain other ways to block design in explaining life. Some theorists think our own little universe is quite enough to render life not only probable, but virtually certain. Stuart Kauffman, for instance, identifies life with "the emergence of self-reproducing systems of catalytic polymers, either peptides, RNA, or others" (*The Origins of Order*, 1993, p.340). Adopting this theoretical perspective, Kauffman develops a mathematical model in which "autocatalytic polymer sets . . . are expected to form spontaneously" (p.288). Kauffman is attempting to lay the foundation for a theory of life's origin in which life is not a lucky accident, but an event that is to be fully expected:

I believe [life] to be an expected, emergent, collective property of complex systems of polymer catalysts. Life, I suggest, 'crystallizes' in a phase transition leading to connected

