American Philosophical Quarterly
Volume 4, Number 2, April 1967

VII. THE PROBABILITY OF THE SIMPLE HYPOTHESIS

GEORGE SCHLESINGER

I

THE current trend among philosophers of science is to treat with suspicion any attempt to justify ontologically the principle of simplicity. Apart from the negative reason that there is no basis for assuming that nature is governed by simple laws, there are two positive arguments to support this attitude: one empirical, the other logical.

First, it is claimed that in the light of the findings of modern science and in the face of its ever-increasing complexity, it is absurd to maintain that nature itself prefers simple laws. We no longer believe, for instance, that the co-variation of the pressure and the volume of a given gas is truly represented by Boyle's simple function, and we know that Kepler's simple laws of planetary motion are rough approximations only, as in fact planets do not move along elliptic orbits but along extremely complicated curves.

The second logical reason is this: Given a number of parameters there is only limited freedom in arbitrarily fixing the relationships among them. Once we have stipulated what form a part of those relationships should take, the form of the rest inevitably imposes itself upon us. Thus if we stipulated that planetary interactions are governed by the simple laws of Newtonian Dynamics, then it inevitably follows that where more than two celestial bodies interact the resulting orbit cannot be a simple elliptic orbit for any one of them. Again, if we stipulate that the simple laws of the Kinetic Theory govern the interplay of molecules, then if the molecules have a finite volume it is easily shown that the resultant macroscopic relationship between pressure and volume cannot be the law enunciated by Boyle. Thus, even if it were the case that nature "preferred" simple laws it would not have the freedom—in a rich universe like ours where there are so many entities and processes, and where there are so many relationships into which physical parameters can enter with one another—to stipulate that all these relationships be simple. Hence the very most a believer in the simplicity of nature could claim is that a small fraction of nature's laws are simple. But then it would seem that the reasonable thing to do, in any individual case, was to expect that the law one is after is, like most laws of nature, a complex one.

The approach favored nowadays, therefore, is to look upon the principle of simplicity as not arising out of any assumptions about the character of nature's laws, but as a principle dictated by some basic methodological propriety. On the surface this may appear a reasonable attitude to take for the objections raised against the ontological justification of the principle of simplicity seem quite convincing as long as one has not too searchingly inquired what exactly is the practical difficulty for the ironing out of which the principle is to be invoked. To get a clear view what kind of basic difficulties one faces when trying to adjudicate among competing hypotheses, we shall consider at some length Jeffreys' views on the subject. His work on simplicity contains invaluable insights.

II

Jeffreys in his *Scientific Inference* asks us to suppose we performed Galileo's inclined plane experiment with the view of obtaining the relationship between the distance s covered by the rolling body and the time t taken to cover that distance. Suppose the following results have been obtained:

t:	0	5	10	15	20	25	30
s:	0	5	20	45	80	125	180

The law which fits these results is given by the equation (a) $5s=t^2$, which is the law enunciated by Galileo. However, the problem arises that there are infinitely many other laws which account equally well for these results.

These laws may be represented by:

(b) $5s= t^2 + t(t-5)(t-10)(t-15)(t-20)(t-25)(t-30)f(t) \ldots$[1]

[1] Sometimes it is erroneously said that (b) does not cover all the possible laws which fit the specified data and that there are others which account equally well for them without conforming to schema (b), e.g.,

Where $f(t)$ can take on infinitely many different forms with the only restriction that it should have finite values for $t = 0, 5, 10, 15, 20, 25, 30$[2]. Why then, does the scientist choose equation (a)?

Of the attempts to explain the scientists' approach as an application of a mere methodological rule he considers only the suggestion that the scientist chooses (a) because it is so much more convenient to handle than (b) and Jeffreys rejects this since he takes it for granted that the scientist sincerely believes that (a) is more likely to work than (b), that is, that for hitherto unobserved values of t, by using equation (a) one will obtain the right prediction and not by using any of the equations of form (b).

Jeffreys then points out that even though any one of the infinitely many laws capable of accounting for the observed results may be true and their probabilities add up to one, this does not entail that the *a priori* probability of each one is zero. For, as he explains, it is possible to have a (denumerably) infinite number of finite terms, the sum of which does not exceed one, as for example is the case with the infinite sequence $1/2 + 1/2^2 + 1/2^3 + 1/2^4 + \ldots$. Thus there is nothing wrong with saying that there are (denumerably) infinitely many possible laws which fit the observed results and each one of these has some non-zero prior probability. For this amounts to no more than claiming that the prior probabilities of all the possible laws are represented by the terms of a continually decreasing convergent series whose sum is one. If in addition we also claim that the simpler the law is, the earlier its probability occurs in the series then we have an adequate explanation why the scientist chooses the simplest law.

It is obvious that the most one can claim for Jeffreys' position is that he had succeeded in showing that if we *had* some good reason to claim that simple laws are more probable than complex laws and that their probabilities varied as the terms of an infinite series whose sum converged to unity, then we could arrange all the admissible laws in the order of their decreasing simplicity and be able to assign a specific finite value to the probability of any one of the infinitely many possible laws. I say this is the most he might have achieved because some have seriously questioned whether indeed he has succeeded in devising a satisfactory method of ordering equations uniquely according to their complexity.[3] Be it as it may, Jeffreys has certainly not provided any proof nor did he claim to have provided a proof that there is no other way of justifying the choice of a given hypothesis on the basis of probability considerations and in a way compatible with the common probability calculus but by assigning the terms of a converging infinite series as the values of the prior probability of hypotheses ordered by their decreasing simplicity.

Obvious as this may strike most of us, it does not appear so to everyone. J. C. Harsanyi who discusses Jeffreys' views on probability and simplicity says, for example, the following:

> The rule that simpler hypotheses should be assigned higher a priori probabilities can be justified in various ways. Jeffreys has shown that this rule directly follows from the axioms of the probability calculus if we add the requirement that no possible hypothesis should be allotted zero probability. For the sum of all the probabilities cannot exceed unity; and this is possible only if listing all the infinitely many admissible hypotheses in order of increasing complexity the probabilities assigned to them form a diminishing number series (possibly with a finite number of exceptions).
>
> There is, however, I believe a more fundamental philosophical reason why simpler hypotheses should be assigned higher a priori probabilities. The reason is that they involve *a smaller number of independent assumptions*.[4]

(c) $5s = t^2 + \sin t . \sin (t - 5) \ldots . \sin (t - 30) \ldots .$

In fact, however, by judicious choice of $f(t)$, (b) reduces to any proposed form. In particular, by letting

$$f(t) = \prod_{n=0}^{6} \left(\frac{\sin (t - 5n)}{t - 5n}\right) \text{ for } t \neq 5k \text{ where } k \text{ is an integer and } 0 \leqslant k \leqslant 6$$

$$\text{and } f(t) = \prod_{n=0}^{k-1} \left(\frac{\sin (t - 5n)}{t - 5n}\right) . \prod_{k+1}^{6} \left(\frac{\sin (t - 5n)}{t - 5n}\right) \text{ for } t = 5k$$

(b) reduces to (c). This point I owe to Drs. A. M. Hasofer and J. C. Burns.

[2] I believe it is fair to say that Jeffreys would agree that $f(t)$ must yet further be restricted so that the resultant variation of s with t conforms to certain commonsense observations known long before Galileo, e.g., that bodies subject to the pull of the earth keep (fairly close) to the same direction throughout their fall (or at any rate throughout the first few hundred yards of their journey).

[3] See R. Harré, "Simplicity as a Criterion of Induction," *Philosophy*, vol. 34 (1959), pp. 229–234; and Robert Ackermann, "A Neglected Proposal Concerning Simplicity," *Philosophy of Science*, vol. 30 (1963), pp. 228–235. The latter shows quite convincingly that Jeffreys has not succeeded to achieve what he set out to do.

[4] J. C. Harsanyi, "Popper's Improbability Criterion for the Choice of Scientific Hypotheses," *Philosophy*, vol. 35 (1960), pp. 309–312.

It is clear, however, that to satisfy Jeffreys' requirements, it is not necessary that we arrange infinitely many admissible hypotheses in order of increasing complexity and assign to each of them a term of a converging series. It is first of all obvious that the same thing can be achieved if we list all the infinitely admissible hypotheses in any other order and assign probabilities to them which correspond to the terms of a convergent number series. In addition the requirement that the sum of all the probabilities should not exceed unity may be reconciled with the fact that no possible hypothesis should be allotted zero probability and furthermore that they are all equiprobable, as long as there are only finitely many admissible hypotheses in any given case. This may be achieved by stipulating that a hypotheses does not qualify as admissible by the mere fact that it satisfies all the experimental results—it also has to meet certain restrictions of form. Then there is yet a third and most important possibility—and this in fact is what I shall attempt to show to be the case—according to which we need not place any implausible restrictions on the number of admissible hypotheses. They may be, as one would naturally assume them to be, non-denumerably infinite. It need not disturb us if the prior probability of any specific hypothesis is zero. What matters is that the "simplicity-conjecture," namely the conjecture that the *kind* of law operating is simple, which initially might have had an extremely low probability, keeps increasing with the increase in the number of experimental data. It also turns out as we shall see these data are compatible but with one of the infinitely many simple laws that were originally admissible, namely with law (a). Consequently the probability that law (a) is true, while prior to experiment it was zero, keeps increasing indefinitely and approaches certainty with the accumulation of experimental results.

It may be of some interest to point out that the claim made in the second paragraph of the quoted passage is not substantiated. There is no reason whatever for saying that to maintain law (b) necessarily requires the making of more assumptions than to maintain law (a). Were we dealing with rival hypotheses which differ from one another in the type of physical mechanism they postulated to be underlying the given phenomenon, Harsanyi's claim might have been relevant. If hypothesis (a) postulated a simple mechanism with a few physical components while hypothesis (b) postulated a complex mechanism in which a large number of factors combined to give rise to the observed phenomenon it would make good sense to say that hypothesis (b) rested on a larger number of independent assumptions than hypothesis (a). But in cases where nothing is said about the mechanism responsible for a given functional relationship and the only question is which mathematical form represents it correctly, it makes no sense to speak about making more independent assumptions by assuming a more complex than a more simple form. In our particular case we either make the single assumption that time and distance vary in accordance with the equation $5s = t^2$, or the alternative single assumption that they vary in a manner represented by the equation $5s = t^2 + t(t-5)$, etc.

III

A solution based on probability considerations can, however, be shown to follow without making any elaborate presuppositions Jeffreys asks us to make. As a first step we must agree that given a finite number of points on a plane, it is not a certainty that a simple curve will pass through them. For my present purposes I regard a curve as simple if and only if the law governing the co-variation of two physical parameters represented by it is regarded as simple by the scientific community. I am not going to give a definition of what is to count a simple law. The point is that no matter what definition we have in mind, or indeed whether we have any definition in mind, we are compelled to agree that there are sets of points, representing parameter couples, through which it is impossible to draw any simple curve. Otherwise we would be forced to say that simple laws and only simple laws operated in any logically possible universe, since given a finite number of observations of the co-variation of any two physical parameters—and we can only be given a finite number of them—a simple law can always account for them.

In reply to this it could perhaps be said that while it would certainly be absurd to assume in all cases of physical parameter-couples that a simple curve may be drawn through a set of points representing their co-variation, we might do so in some selected instances. If so, all we have to ask is whether the co-variation of time and distance of freely falling bodies belongs to this select group? One who answers *Yes*, need go no further; for him the use of the principle to choose the simple law compatible with the data rather than any of its complex rivals is fully justified. But not many shall be happy with such a justification. That is why the problem we set out to solve will seem to most a very real one, namely, how to justify the use of the principle of simplicity in a situation like the one confronting Galileo? Thus we take it that the answer is *No*: We do not assume that a simple curve may be drawn through any number of points representing the co-variation of s and t where the values of one of them has been chosen at random. Having been granted this, we shall also agree then, of course, that if n is the minimum number of such points through which it is no longer a certainty that a simple curve may be drawn, then as we increase the number of those points beyond n, the probability that a simple curve will pass through them keeps decreasing. The only additional assumption we shall use in what follows, and this will be granted at once, is that *some* curve will certainly pass through a finite set of points no matter how large.

Now let C stand for the proposition: A complex curve passes through the points on the paper in front of me.

And let S stand for the proposition: A simple curve passes through the points on the paper in front of me.

And let L stand for the proposition: The law governing the co-variation of time and distance belongs to that *large* group of laws of nature which are represented by a *complex* curve.

And let T stand for the proposition: The law governing the co-variation of time and distance belongs to that particular group of laws of nature which are represented by a *simple* curve.

And let P_n stand for the proposition: There are n points on the paper in front of me and they represent the results of experiment to determine the co-variation of time and distance.

Employing the usual notation where "$p(T/P_nS)$" stands for "the probability of T given that P_n and S" we have by the conjunctive axiom of probability

$$p(T \,.\, S/P_n) = p(T/P_n) \,.\, p(S/T \,.\, P_n) = p(S/P_n) \,.\, p(T/S \,.\, P_n).$$

$$\text{Hence } p(T/S \,.\, P_n) = \frac{p(T/P_n) \,.\, p(S/T \,.\, P_n)}{p(S/P_n)}$$

But since P_n and T together entail S, $p(S/T \,.\, P_n) = 1$. (If it be objected that we have disregarded the possibility of experimental error let S stand for: a simple curve passes through or *sufficiently close to* each one of the points on paper in front of me.)

$$\text{Thus } p(T/S \,.\, P_n) = \frac{p(T/P_n)}{p(S/P_n)}$$

Employing the disjunctive and then the conjunctive axiom of probability we get:

$$\begin{aligned} p(S/P_n) &= p(T \,.\, S/P_n) + p(L \,.\, S/P_n) \\ &= p(T/P_n) \,.\, p(S/T \,.\, P_n) + p(L/P_n) \,.\, p(S/L \,.\, P_n). \end{aligned}$$

Now, since as we have said $p(S/T \,.\, P_n) = 1$, the first term of the R.H.S. reduces to $p(T/P_n)$. The prior probability of L/P_n may be assumed as near unity as we please. The crucial point is that $p(S/L \,.\, P_n)$ keeps decreasing with increasing n as explained. Thus the value of $p(S/P_n)$ keeps approaching $p(T/P_n)$ from which it follows that $p(T/S \,.\, P_n)$ gets closer and closer to unity with the increase of n. We see therefore that while the prior probability that the law we are after belongs to the small set of laws characterized by their simplicity may not be appreciable, it greatly increases after obtaining experimental results and seeing that a simple curve passes through the points representing them. Indeed the conjecture that T, given that S and P_n approaches certainty. But eventually the only law left which is both compatible with T and the experimental results is law (a);[5] hence the probability that law (a) is the true law varies as $p(T/P_n \,.\, S)$.

On the other hand the probability that out of the infinitely many complex equations of form (b) any particular one represents the true law of nature is zero even initially when the probability of L may be taken as near to unity as we please (unless we assumed that the admissible complex equations were merely denumerably infinite and that their probabilities varied as the terms of a converging series). Subsequently, of course, the probability of L itself decreases indefinitely as the probability that the relationship between time and distance is truly represented by $5s = t^2$ steadily increases with our collection of data.

IV

All this may strike us, however, as too good to be true and we may begin to think that we achieved what we did by dividing all curves into two groups

[5] For a detailed argument supporting this contention see the latter part of this article.

just in such manner as to ensure the desired result. Thus the suspicion arises that one could devise an argument which would just as powerfully show that the probability of some particular law represented by a specific equation of form (b) keeps increasing with the accumulation of experimental results on the basis of the division of all curves under a different principle.

Let us have this objection laid out in detail: Suppose we compile a small list of curves which will include one particular curve of form (b) and a few dozen other curves chosen in any way one may think of—with the only restriction that given a finite number of points on a plane it should not always be a certainty that some curve, the kind of which is on our list, passes through them. Let the curve represented by (a) not be on this small arbitrary list of curves. Now let :

C' stand for the proposition: A curve *not* to be found on our small list of arbitrarily selected curves passes through the points on the paper in front of me.

S' stand for the proposition: A curve *to be* found on our small list of arbitrarily selected curves passes through the points on the paper in front of me.

L' stand for the proposition: The law governing the co-variation of time and distance belongs to that large group of laws which are represented by a curve *not* to be found on our small list of arbitrarily selected curves.

T' stand for the proposition: The law governing the co-variation of time and distance belongs to that particular group of laws which are represented by a curve *to be* found on our small list of arbitrarily selected curves;

then it appears that $p(T'/P_n \,.\, S)$ approaches one for exactly the same reason as $p(T/P_n \,.\, S)$ approaches one and this completely destroys our previous justification for choosing (a).

The crucial point, however, which has been overlooked in this argument is that even though $p(S'/P_n)$ may be approaching $p(T'/P_n)$, dividing $p(T'/P_n)$ by $p(S'/P_n)$ does not necessarily give us a large number or indeed finite number, if to begin with $p(T'/P_n)$ equals zero! And the scientist will in fact assign zero credibility to T'. After all, there are infinitely many ways in which arbitrary lists may be compiled, and, as one list is as good as the other, it is not unreasonable for the scientist to assign zero probability to the representability of the law we are after, by a curve belonging to a specific arbitrary list. However, the "list" which contains all the simple curves enjoys a slightly privileged status. It is not required that the scientist should believe that all laws are simple, nor even need he be convinced that some laws are simple. It is sufficient that he maintains that there is a finite probability that some of the laws which are likely to come to our attention are simple: in other words the initial probability that T, however small it may be, is not zero. The question what the precise status of this assumption is will be discussed in Section VI.

V

Let me now make explicit the crucial points I hope to have established if the approach outlined in the previous section works.

First, I have been anxious to ensure that it is understood that the strong belief in a typical post-experimental situation that the simple law is immensely more likely to be true than any one of its infinitely many rivals is ultimately based on a far more innocuous assumption than it appeared at first sight—an assumption not involving the belief in the considerable prior likelihood of a simple law operating in a given case. An important aspect of the situation on which I have tried to focus attention is, however, that it is in the nature of things as elementary mathematics show, that our modest initial assumption, of necessity, becomes immensely magnified upon finding that the large number of experimental results obtained, all fit—within a reasonable margin of experimental error—a simple law.

Secondly, it should be clear now why we need not worry about devising a generally applicable criterion restricting the number of admissible laws prior to experiment to ensure the effective functioning of the principle of simplicity, nor, of course, about the defendibility of the criterion chosen. The number of laws which may be regarded as candidates for being operative in any given case need not be restricted prior to experiment. It should not disturb us if their number is an infinite one and non-denumerably so. Our argument is not based on the finite prior probability of *any given law*. It is based on the finite prior probability of T and L—e.g., on the finite probability that the law operative in our case belongs to this or that *kind* of law.

Thirdly, and most importantly, the vexing, if not insoluble problem of how to give a precise definition of what makes a function simpler than another and how to find an adequate method whereby all the different admissible functions may be graded according to their increasing complexity should now be seen as irrelevant to the really significant problem of how the principle of simplicity works. Surprising as it may sound at first, but according to the present explication, the principle achieves its purpose even though we entirely ignore the question of the relative simplicity of functions and make no more than some very general statements as to what makes a function categorically complex. A function may be said to be definitely non-simple if it is too complex to be handled usefully in the context of the conceptual methods available at a given time and only if it has never in the course of the history of science been advanced as representing a true law of nature.

To demonstrate that this is indeed so, let me first answer what on the surface may seem a serious charge against my account, namely, that I have failed to say anything about a situation where a number of simple hypotheses compete against one another. Suppose a variety of hypotheses—each regarded by the scientific community as simple to one degree or another—may account for all the results obtained. We have neither provided any rule as to how to distinguish between their varying degrees of simplicity, nor a basis for justifying our choice of the simplest among these.

What we must realize, however, is that such a situation is, from a philosophical point of view, relatively uninteresting. All we have to do to escape our predicament is to perform some further experiments for such values of the parameters involved, with respect to which the various rival hypotheses yield different predictions. Whatever the results, the wrong hypotheses will then be eliminated. And there must be such values, otherwise the hypotheses would be equivalent and not competing with one another. In the context of the profound and philosophically really important problem to which Jeffreys has drawn our attention, the difficulty is an ineradicable one. Although each additional experiment eliminates an infinite number of hypotheses there are always still infinitely many of them left in the running.

Suppose that in the context of Galileo's experiment we perform additional experiments for times $t_1, t_2, \ldots, t_n$ and we find the results all fitting equation (a). We would then have eliminated all those (b) type equations in which $f(t)$ does not contain the term $(t-t_1)\ (t-t_2)\ \ldots\ (t-t_n)$, but all the other equations in which $f(t)$ does contain this term, and the number of such (b) type equations is infinite, remain to compete against (a). There is no escape from this predicament but by the use of the argument that (a), even though it is only one of infinitely many equations which are equally well satisfied by the experimental data, is immensely more likely to represent the true law operating than any of its rivals for the reasons elaborated here.

Now it should be possible to see clearly why we need nothing more precise than the kind of statement I made in attempting to describe what constitutes a categorically complex function. That statement, to be sure, does not provide a secure enough basis upon which to determine with certainty, whether in the situation described by Jeffreys, we are to regard equation (a) as the sole simple function compatible with the experimental data or perhaps some functions of type (b) may too be so regarded. This, however, need not worry us very much. Some may feel that (a) is the only function which is not categorically complex. According to them, good reason has been provided for preferring (a) to any of the other infinitely many functions.

Suppose, however, that there are others who think that some of the equations of form (b) with a simple kind of $f(t)$ are also not to be regarded as categorically complex. They indeed would not have any good reason to adopt a particular function as the one representing the law of terrestial free fall *at this stage of the experiment*. What they would have to do, therefore, is to obtain some more experimental data to eliminate more hypotheses until a stage is reached where only one of the surviving functions is simple and the remainder categorically complex.

It is of course not a certainty that such stage will ever be reached, or if reached, it will be permanently maintained. It may well happen that all simple functions become casualties of further experimentation. In that case, this particular line of investigation will have to be abandoned. Nothing may be hoped from continuing to observe the covariation of time and distance since there is no longer any basis upon which to adjudicate among the competing hypotheses. The law of the covariation of time and distance may then be hoped to become disclosed only as a consequence of other laws simple enough to be amenable to direct investigation and discovery.

VI

According to the present suggestion, then, the ontological justification of the principle of simplicity need not be based on anything more than the claim that the initial probability in any given case that the law we are after is not zero. It is still a pertinent question which one is entitled to ask: What is the justification of this claim?

First, we should take notice of the fact that the positive arguments mentioned in Section I do not apply against the present justification. That there is no empirical evidence and could be no empirical evidence against the claim on which we based our justification is only too evident, since all we insist upon now before we apply the term "simple" to a function, that its kind has (because of its forbidding complexity) never throughout the history of science been advanced as representing a true law of nature. But it must be obvious also that the second, logical argument does not apply either. We may readily admit that prior to experiment the probability that any given functional relationship is a very complex one is vastly greater than that it is manageably simple. This, however, does not interfere with the proper functioning of the principle which requires no more than that T should not equal zero initially for then it keeps increasing indefinitely with the number of experiments it successfully survives.

Some may wish to go further and maintain that not only are there no positive arguments against assigning non-zero probability to T, but that there are arguments why we should definitely not assign zero probability to it. A believer in the intelligibility of nature, according to whom at least some, if not all, the laws of nature are discoverable by humans, must postulate that finite prior probability attaches to the conjecture that a given law belongs to *some* specific kind, otherwise, as we have seen, no amount of evidence can provide the slightest support to any functional law. The only question then is which particular kind of function be accorded preferential status. Believing in the intelligibility of nature entails the belief that among the laws held at the moment to be true, there is a considerable proportion which is actually true. If so, it seems reasonable to assign finite probability to the event that the law presently under investigation is of the same kind as those already known to us. But the one thing one could safely say about all the laws already known to us to have in common is that they are not categorically complex in the sense earlier defined. As long as no one suggests some other feature that is common to all the laws hitherto entertained, finite probability attached to the conjecture that the desired function is a member of the class of functions characterized by their simplicity, but not to the conjecture that it is a member of some other class.

Others who subscribe to the principle of the intelligibility of nature may wish to go further and argue that a finite proportion of all the laws operating in the universe must be simple for if this were not so, the resulting phenomena would be of such frightful complexity that no simple regularities would be anywhere apparent to provide any foothold upon which to begin a systematic and rational investigation of nature. Yet others would be prepared to concede that the group of laws which are simple may very likely amount to a vanishingly minute fraction of all the laws which govern the universe. The reason why we nevertheless do not assign vanishingly small initial probability to the assumption that a given law is simple is that the very simplicity of a law lends it a unique advantage which raises the likelihood of its discovery to a level which is out of proportion with the distribution of simple laws in nature. A simple law attracts more readily the attention and secures more easily the comprehension of the scientist.

It is conceivable, however, that valid objections can be raised against all these arguments. Some might be of the opinion that in general any attempt to justify an assumption of such fundamental nature as the one that the initial probability of T is not zero must of necessity fail. It must be recognized that ultimately science rests on a number of unprovable assumptions. What one may demand with respect to such assumptions is that it be shown that without them rational inquiry is impossible; that they be kept at a minimum and yet be able to do the job expected from them; that they should not be counter-intuitive and that there be no other positive arguments against their credibility. The assumption that nature is simple in the sense here explicated seems to satisfy these requirements.

Received April 14, 1966

The Australian National University
and
The Minnesota Center for the Philosophy of Science