A Bayesian Analysis of the Design Argument

A Dissertation submitted in partial fulfilment of the requirements for the degree of Master of Philosophy,

by,

Romain Rabot

October 1981
OUTLINE

(A) Acknowledgements

(B) Summary

(C) Introduction

(D) The Confirmation of Scientific Hypotheses
   (i) The hypothetico-deductive method
   (ii) Hanson's logic of discovery
   (iii) Popper's method of corroboration
   (iv) Causal hypotheses and Bayes's theorem
   (v) The status of prior probabilities
   (vi) Criteria for plausibility judgments

(E) The Intelligent Design Hypothesis

(F) Conclusion

(G) Bibliography
I wish to thank Dr Anthony O'Hear for his advice and encouragement, and my wife, Sue, for her patient typing and retyping of numerous drafts of this essay.
SUMMARY

Theists have claimed that the existence of God can be approached as a scientific hypothesis, and that His existence can be established with a high degree of confirmation by observational evidence. The design argument proceeds from "the curious adapting of means to end" throughout all nature which resembles "the productions of human contrivance" to an "Author of Nature" who being "somewhat similar to the mind of men" is responsible for all these phenomena.

Arguments by analogy, which are much more subtle and complex than simple analogies, occur also in certain areas of scientific research and have been classified as an important type of inductive argument. Their function is to evaluate causal hypotheses. So, they are best understood when subjected to analysis appropriate to arguments offered in support of causal hypotheses in science.

Analysis of the inference by which scientific hypotheses are confirmed by observational evidence shows that its structure is given by Bayes's theorem. This schema could be used to turn the hypothetico-deductive method into a valid one by supplementing it by two kinds of probabilities. First, we must assess the probability that our observational results would obtain even if the hypothesis under consideration were false. The smaller the probability the stronger the degree of confirmation. This seems a natural interpretation of Popper's methodological requirement that scientific hypotheses must be audacious and take risks. We must also assess the prior probability of the hypothesis we are considering. This
is a reasonable interpretation of Hanson's demand for plausibility arguments. The probabilities that enter into this Bayesian schema are interpreted as frequencies. This enables us to show the relevance of probabilities to prediction, theory and practical decision.

On the Bayesian analysis, the design argument, far from supporting the existence of God, provides strong evidence to the contrary.
INTRODUCTION

In both pagan and Christian traditions, the occurrence of order in the natural world is the most widely used and generally accepted ground for arguing from the world to the existence of an intelligent and powerful creator. The argument, in various forms, is to be found in Plato, Xenophon and Cicero; in Aquinas, Newton and Berkley; and in a great many of the eighteenth century attempts to establish the reasonableness of religion. In the nineteenth century William Paley's version enjoyed great popularity and it is in his Natural Theology that the classic statement of this argument is found.

"In crossing a heath, suppose I pitched my foot against a stone, and were asked how the stone came to be there, I might possibly answer, that, for anything I knew to the contrary, it had lain there forever: nor would it perhaps be very easy to show the absurdity of this answer. But suppose I had found a watch upon the ground, and it should be inquired how the watch happened to be in that place, I should hardly think of the answer which I had before given, that for anything I knew, the watch might have always been there. Yet why should not this answer serve for the watch, as well as for the stone? Why is it not as admissible in the second case, as in the first? For this reason, and no other, viz. that, when we come to inspect the watch, we perceive (what we could not discover in the stone) that its several parts are framed and put together for a purpose, e.g., that they are so formed and adjusted as to produce motion, and that motion is so regulated as to point out the hour of the day; that, if the several parts had been differently shaped from what they are, of a different size from what they are, or placed after any other manner, or in any other order, than that in which they are placed, either no motion at all would have been carried on in the machine, or none which would have answered the use, that is now served by it" (Paley, Natural Theology, p.5).

It is the same argument that Hume puts in the mouth of Cleanthes:
"I shall briefly explain how I conceive this matter. Look round the world: contemplate the whole and every part of it: you will find it to be nothing but one great machine, subdivided into an infinite number of lesser machines, which again admit of subdivisions, to a degree beyond what human senses and faculties can trace and explain. All these various machines and even their most minute parts, are adjusted to each other with an accuracy, which ravishes into admiration all men, who have ever contemplated them. The curious adapting of means to ends through all nature, resembles exactly, though it much exceeds, the productions of human contrivance of human design, thought, wisdom, and intelligence. Since therefore the effects resemble each other, we are led to infer, by all the rules of analogy, that the causes also resemble; and that the Author of Nature is somewhat similar to the mind of men; though possessed of much larger faculties proportioned to the grandeur of the work, which he has executed. By this argument a posteriori, and by this argument alone, do we prove at once the existence of a Deity, and his similarity to human mind and intelligence" (Hume, *Dialogues Concerning Natural Religion*, p. 22).

The design argument is clearly an argument from analogy in that it appeals to observed similarities between human artifacts and works of nature. It must also, I think, be classified as empirical.

Arguments from analogy have been classified as an important type of inductive argument, and occur in certain areas of scientific research. Important applications of analogical reasoning are found in the field of medicine where experiments are carried out on animals in order to ascertain the effects of various substances upon humans. If we characterized such arguments as merely analogical we would run the risk of seriously misrepresenting them and overlooking much of their force. The strength of a simple argument by analogy depends crucially upon the degree of similarity between the entities with respect to which the analogy is drawn. There is one obvious sense, however, in which such arguments are not simple analogies. This is that it is not the degree of similarity that matters most, but the relevance of the similarities and the irrelevance of the dissimilarities.
Such arguments initially characterized as analogies are really more subtle and complex. They are arguments whose function is to evaluate causal hypotheses. If we are to understand them we must subject them to the sort of analysis appropriate to arguments offered in support of causal hypotheses in science.
Science is concerned not only with matters of discovery but also with matters of justification. The history of science contains many instances of old theories being replaced by new ones. But on what bases were the replacements made? One's interest is aroused in such questions as whether, to what extent, and in what manner the old theory was disconfirmed; and similarly, what evidence was offered in support of the new hypothesis, and how adequate such evidence was. One might even ask whether nonevidential factors such as national rivalry among scientists, and esthetic disgust with certain types of theories operate in the selection of alternate hypotheses.

The most widely held picture of scientific confirmation is the hypothetico-deductive method. According to this view, a scientific hypothesis is tested by deducing observational consequences from it, and seeing whether these consequences transpire. Some empirically determined initial conditions are supplied to make it possible to validly deduce the observational consequences; and auxiliary hypotheses are frequently made use of to connect the observations with the hypothesis that is being tested. The hypothetico-deductive method can be schematized as follows:

\[ H \] (hypothesis being tested)
\[ A \] (auxiliary hypotheses)
\[ I \] (initial conditions)
\[ \ldots O \] (observational consequence)
Let us assume that the initial conditions I have established as true by observation, and that we can ascertain by observation whether the observational consequence 0 is true or false. Let us assume, moreover, that for purposes of the present test of our hypothesis H, that the auxiliary hypotheses A are unproblematic. Having made these simplifying idealizations, we can say that H implies 0. If 0 turns out to be false, it follows that H must be false. This is deductively valid *modus tollens*. Given the truth of 0, however, nothing follows deductively about the truth of H. We would commit the deductive fallacy of affirming the consequent should we infer the truth of H from the truth of 0. According to the hypothetico-deductive view, the truth of 0 does tend to confirm or lend probability to H. If enough observational consequences are bourne out by experience the hypothesis can become quite highly confirmed. Scientific hypotheses can never be completely and irrefutably verified in this way, but they can become sufficiently confirmed to be scientifically acceptable. According to this hypothetico-deductive conception induction is a kind of inverse of deduction. The fact that a true observational prediction follows deductively from a given hypothesis means, according to the hypothetico-deductive view, that a relation of inductive support runs in the reverse direction from 0 to H.

This account is inadequate. We need a more satisfactory account of scientific confirmation. The main shortcoming of the hypothetico-deductive method is that it gives us no basis whatever on which to choose between alternative hypotheses which are confirmed in exactly the same manner. It clearly stands in need of supplementation.

**Hanson's Logic of Discovery:**

We must take care in posing the question of the existence of a
logic of discovery for empirical science. It would be fantastic to suggest that there might be a mechanical method that necessarily generates true explanatory hypotheses. Problems of discovery aside, there is no way of determining that we have a true hypothesis. The logic of science cannot be regarded as an algorithm that would yield all scientific truth. But what could be reasonably demanded of a logic of discovery should such a thing exist? Hanson, and Peirce before him, answer not that it must generate true hypotheses but plausible conjectures. Hanson believed that this demand could be fulfilled. He began by distinguishing between "reasons for accepting a hypothesis \( H \)" and "reasons for suggesting \( H \) in the first place". Reasons for accepting \( H \) will be reasons we have for thinking \( H \) true. But the reasons for suggesting \( H \) originally, or for formulating \( H \) in one way rather than another, are those reasons which make \( H \) a plausible type of conjecture. Hanson grants that non-logical factors play a part in the discovery of hypotheses, but he claims that there are, in addition, perfectly good logical reasons for regarding hypotheses of a particular kind as being those most likely to succeed. These reasons are logically distinct from the kinds of reasons that later, in the case of successful hypotheses, make us elevate hypotheses from the status of plausible conjectures to the status of acceptable, true, or highly confirmed hypotheses. Salmon distinguishes three logically distinct aspects of the treatment of scientific hypotheses. Those can be taken in any order, and can, he says, be mixed together. There are still three logically distinct matters: (i) thinking of the hypothesis, (ii) plausibility considerations, and (iii) testing or confirmation. Salmon thinks that Hanson has correctly distinguished between plausibility arguments and the testing of hypotheses, but mistakenly conflated plausibility
arguments with discovery. Using Kepler as an example, Hanson discusses hypotheses that would have been rejected by Kepler as implausible.

"Other kinds of hypotheses were available to Kepler: for example, that Mars' color is responsible for its high velocities, or that the dispositions of Jupiter's moons are responsible. But these would not have struck Kepler as capable of explaining such surprising phenomena. Indeed, he would have thought it unreasonable to develop such hypotheses at all, and would have argued thus. (Hanson, "Is There a Logic of Discovery?" Current Issues in the Philosophy of Science. p.23)

Kepler would no doubt have rejected such hypotheses had they occurred to him. But thinks Salmon, there is no reason to suppose that these considerations were psychologically efficacious in preventing Kepler from thinking of such hypotheses – they might, however, have been efficacious in preventing him from mentioning them - and in causing him to think of others instead. Furthermore, he suggests, it does not matter in the slightest. What matters is that, had such unreasonable hypotheses crossed Kepler's mind, plausibility arguments alone would have sufficed to prevent them from coming to serious empirical testing.

One basic question remains. Plausibility arguments can be distinguished from hypothesis testing and confirmation on the one hand and from the psychology of discovery on the other. But, what precisely is their status? Are plausibility considerations psychological or subjective in character? Do they play a legitimate role in science or do they merely reflect the prejudices of the scientific community? Are they different in kind from the considerations involved in the confirmation of hypotheses? An answer will be forthcoming when we consider what the probability calculus tells us about confirmation.
Popper's Method of Corroboration

Karl Popper's deductivism is one of the most interesting and controversial attempts to provide an account of the logic of science. The hypothetico-deductive method being ampliative and non-demonstrative, is not strictly deductive, and for as long as it is regarded as a method of supporting scientific hypotheses, cannot succeed in making science thoroughly deductive. Popper realizes this, and in arguing that deduction is the sole mode of inference in science rejects the hypothetico-deductive method as a means of confirming scientific hypotheses. He denies that induction plays any role whatever in science. Indeed, he maintains that there is no such thing as a correct inductive argument. He admits the psychological fact that people have faith in the uniformity of nature, but this, he says, is no more than a matter of psychological fact.

Popper's fundamental thesis is that it is falsifiability that characterizes statements of empirical science and distinguishes them from metaphysical statements and from tautologies. Scientific hypotheses are general in form and so are amenable to falsification but not verification. One negative instance suffices to falsify a universal generalization; and no limited number of positive instances can verify a universal generalization. Positive instances are completely indecisive. It is the aim of empirical science to devise theories to stand the test of every possible serious attempt at falsification. Scientific theories are conjectures or hypotheses; they are general statements whose purpose is to explain the world and make it intelligible. They are not to be regarded as final truths. Their status can never be more than that of tentative conjecture. They must always face the severest possible criticism. It is for the theoretician to propose scientific conjectures and the experimentalist to devise every possible way of falsifying these
theoretical hypotheses. The attempt to confirm hypotheses is no part of the aim of science.

General hypotheses in conjunction with statements of initial conditions entail predictions of particular events. We have a very high degree of intersubjective agreement concerning the initial conditions. We can likewise obtain intersubjective agreement as to whether the predicted fact occurred. Should the predicted fact fail to occur, the theory is considered to have suffered falsification. The mode of inference is deduction. If the theory were true, then, given the truth of the statements of initial conditions, the prediction would have to be true. The prediction as it happens, is false; therefore, the theory is false. This is the familiar principle of modus tollens. According to Popper this is the only kind of inference applicable to the acceptance or rejection of scientific hypotheses. It is, however, clearly suitable for rejection only.

Hypothetico-deductive theorists would maintain that they had a confirming instance for the theory if the predicted event occurs. Confirming instances enhance the probability of the hypotheses. With enough confirming instances the probability of the hypothesis becomes great enough to warrant provisional acceptance. With sufficient inductive evidence of this kind we are justified in regarding it as well established. Popper rejects this positive account. According to him if a hypothesis is tested and the result is negative, we can reject it. But on the other hand, if the test is positive all we can say is that we have failed to falsify it. There is nothing unique about a hypothesis that survives without being falsified. It might be that many other unfalsified hypotheses remain to explain the same facts. So, if science is to amount to more than a mere collection of our observations and their various reformulations, it
must embody some other methods besides observation and deduction.

Popper supplies that additional factor: corroboration.

A falsified hypothesis is replaced by one which has not yet been falsified. Not all unfalsified hypotheses are on a par. It is possible to select from among unfalsified hypotheses. Again, falsifiability is the key. Hypotheses differ from one another with respect to the ease with which they can be falsified. And so they can often be compared with respect to degree of falsifiability.

Science, Popper maintains, is interested in bold conjectures. These must be consistent with known facts, but must also run as great a risk as possible of being controverted by facts still to be accumulated. The search for additional facts should be guided by the effort to find facts that will falsify the hypothesis.

As Popper characterizes falsifiability, the greater the degree of falsifiability of a hypothesis, the greater its content. Tatologies lack empirical content because they are compatible with any possible world. Empirical statements are not compatible with every possible state of affairs. The greater the number of possible states of affairs excluded by a statement, the greater its content, for the more it does to pin down our actual world by ruling out possible but non-actual states of affairs. At the same time, the greater the range of facts ruled out by a statement the greater the risk it runs of being false. A statement of high content is more vulnerable to falsification than one of low content because it has more potential falsifiers. So high content means high falsifiability. At the same time, content varies inversely with probability. The logical probability of a hypothesis is defined in terms of its range, or, in other words, the possible state of affairs with which it is compatible.

The greater the logical probability of a hypothesis, the fewer are
its potential falsifiers. Thus, high probability means low falsifiability. Popper's choice is the opposite of that of hypothetico-deductive theorists who would recommend the selection of the most probable hypothesis from among those hypotheses that are compatible with the available facts. He recommends selecting a hypothesis of low probability. Highly falsifiable hypotheses which are severely tested become highly corroborated. The greater the severity of the tests the greater the corroboration of the hypothesis that survives them. Hypotheses are not regarded as true because they are highly corroborated. Increasing corroboration is not a process of accumulating positive instances to increase the probability of the hypothesis.

The crucial difference between Popper's deductivism and the hypothetico-deductive theorists is that the latter attempt to start with probable hypotheses and try to find further support for them through positive confirmations. When more than one hypothesis is available to explain all the available data, the hypothetico-deductivist would choose the most probable one. Popper, by contrast, would begin with the least probable hypothesis, for probability is related inversely to content. Popper would seek to falsify the hypothesis. Failure to do so tends to increase the degree of corroboration. When more than one hypothesis remains unfalsified, we select the least probable one.

Salmon makes the following point about Popper's theory. In furnishing a method for selecting hypotheses Popper exceeds the content of the relevant available basic statements. Valid deductions are nonampliative and their conclusions cannot exceed their premises in content. So demonstrative inference alone cannot accomplish the task of selecting from among unflasified hypotheses. Popper does not pretend that basic statements plus deduction can give us
scientific theory; instead he introduces corroboration. Corroboration is a nondemonstrative form of inference. It is a way of providing for the acceptance of hypotheses even though the content of these hypotheses goes beyond the content of the basic statements. Modus tollens without corroboration is empty, modus tollens with corroboration is induction. So Popper's method must be characterized as a form of inductivism.

Hanson attacks the hypothetico-deductive method for failure to take into account plausibility arguments. Popper attacks the same method for failure to incorporate implausibility considerations. While Hanson's and Popper's attacks are mutually incompatible, each has a valid foundation, and each points to a fundamental shortcoming of the hypothetico-deductive method. It seems that the logical gaps in the hypothetico-deductive method can be filled by means of the ideas suggested by Hanson and Popper and that these ideas lead us to indispensable aspects of the logic of scientific inference.
Causal Hypotheses and Bayes' Theorem.

The basic problem with hypothetico-deductive inference is that it leaves us with a superabundance of hypotheses, all equally adequate to the available data. Each is confirmed in precisely the same manner by the same evidence. A hypothesis is said to be confirmed when, in conjunction with true statements of initial conditions, it entails a true prediction. Any number of other hypotheses that, in conjunction with the same or different statements of initial conditions, entail the same prediction are confirmed in the same way by the same evidence. It is always possible to construct an unlimited supply of hypotheses to fit the bill. The hypothetico-deductive method is, therefore, hopelessly inadequate with respect to determining the acceptability of scientific hypotheses on the basis of empirical data.

When we look around for a more adequate account of scientific confirmation, it is natural to see whether the mathematical calculus of probability can offer any resources. If we claim that the process of confirmation is one of lending probability to a hypothesis in the light of evidence, it is reasonable to see whether there are any theorems on probability that characterize confirmation. If so, such a theorem would provide some sort of valid schema for formal confirmation relations. Bayes's theorem seems well suited for this role. Salmon uses the following example to illustrate its application.

Suppose that a small percentage of pearls have a particular sort of colour flaw which renders them worthless. This flaw appears in 1% of all cultured pearls, and in 3% of all natural pearls. Assume that 90% of all pearls examined are cultured pearls. Now, a pearl is found which exhibits this undesirable colour flaw. What is the probability that it is a cultured pearl?
Bayes's theorem can be used to solve this problem. It reads:

\[ P(A \cap C \mid B) = \frac{P(A \cap B) \cdot P(C \mid A \cap B)}{P(A \cap B) \cdot P(B \mid A) \cdot P(C \mid B) + P(A \cap B) \cdot P(B \mid A) \cdot P(C \mid B)} \]

Let \( A \) be the class of pearls, \( B \) the class of cultured pearls, \( \overline{B} \) the class of natural pearls, and \( C \) the class of colour-flawed pearls. The formula involves the following probability expressions:

- \( P(A, B) = \) the probability that a pearl is cultured = 0.9
- \( P(A, \overline{B}) = \) the probability that a pearl is natural = 0.1
- \( P(A, B, C) = \) the probability that a cultured pearl is colour-flawed = 0.01
- \( P(A, \overline{B}, C) = \) the probability that a natural pearl is colour-flawed = 0.03

These two probabilities are called prior probabilities.

It is important to note that they are not probabilities of hypotheses, but, rather, probabilities of the effect. It is the posterior probability that we seek when we wish to determine the probability of the hypothesis in terms of the given evidence.

\( P(A \cap C, B) = \) the probability that a colour-flawed pearl is cultured. This is the posterior probability. Substituting the values stipulated we find that the posterior probability

\[ P(A \cap C, B) = \frac{0.9 \times 0.01}{0.9 \times 0.01 + 0.1 \times 0.03} = \frac{3}{4} \]

Notice that, although \( P(A, \overline{B}, C) \geq P(A, B, C) \), \( P(A \cap C, B) \geq P(A, \overline{C}, B) \). The inverse probabilities invert the order of the likelihoods.

In order to apply Bayes's theorem, we need the three probabilities \( P(A, B) \), \( P(A, B, C) \), and \( P(A, \overline{B}, C) \). Since the two prior probabilities must add up to one, it is sufficient to know one of them, but the
likelihoods are independent, so we must have both of them. Thus, in order to compute the posterior probability of a hypothesis, we need (i) its prior probability, (ii) the probability that we would get the evidence we have if it is true, and (iii) the probability that we would get the evidence we have if it were false. None of these three is dispensable, except in a few obvious special cases such as if \( P(A,B) = 0 \) or \( P(A,B,C) = 0 \); if \( P(A,B) = 0 \) or \( P(A,B,C) = 0 \), then \( P(A,C,B) = 1 \).

In the hypothetico-deductive method, it is stipulated that the hypothesis being tested implies the evidence. So, in that case \( P(A,B,C) = 1 \). This value of one of the likelihoods does not determine a value for the posterior probability and, indeed, the posterior probability can be arbitrarily small even in the case supplied by the hypothetico-deductive method. This fact exposes the inadequacy of the hypothetico-deductive schema quite dramatically. Even though the data confirm the hypothesis according to the hypothetico-deductive view, the posterior probability of the hypothesis in the light of the available evidence may be even as small as zero in the limiting special case in which the prior probability of the hypothesis is zero.

The hypothetico-deductive method is fallacious as it stands, but can be rectified by supplementing it with the remaining elements required by Bayes’s theorem. According to the hypothetico-deductive schema an hypothesis \( H \) and statements of initial condition (which we will assume to be true and unproblematic) \( I \), an observational prediction \( O \) is deducible. \( H \) implies, \( O \). We have not as yet discussed all the problems involved in interpreting Bayes’s theorem. So for the present, we can provide a loose and preliminary interpretation of Bayes’s theorem. Let "A" refer to hypothesis like \( H \), "B" to the property of truth; and "C" to the observed result with respect to
the prediction 0. If positive confirmation occurs "C" means that 0 obtain; in the negative case "C" designates the falsity of 0. This interpretation makes the expression on the left hand side of Bayes's theorem refer to precisely the sort of probability we are interested in - i.e. "P(A.C,B) designates the probability that a hypothesis of the sort in question, for which the given observational results have been found, is true. This is the probability we seek in dealing with the confirmation of scientific hypotheses. We have seen that we need three probabilities in order to compute the posterior probability of our hypothesis. The hypothetico-deductive method provides only one of them. Given that H implies 0 and that 0 obtains, P(A.B,0) = 1. Bayes's theorem, however, reveals that this value is entirely compatible with a small posterior probability for the hypothesis. A small value for P(A,0) and a large value for P(A.B,0) nullify any tendency of the confirmation to enhance the value of P(A.C,B). Successful confirmation requires all three probabilities. Only one of these is provided by the hypothetico-deductive method. Bayes's theorem embodies the asymmetry between confirmation and falsification. If H implies 0 and 0 does not obtain, then P(A.B,0) = 0, and it follows that the posterior probability of the hypothesis, P(A.C,B), likewise equals zero. Falsification holds a special place in the logic of scientific inference, as Popper has emphasized.

If Bayes's theorem provides a correct formal schema for the logic of confirmation and disconfirmation of scientific hypotheses, it indicates that we need to take into consideration three factors in attempting to assess the degree to which a hypothesis is rendered probable by the evidence. Roughly, it says that we must consider how well our hypothesis explains the evidence we have, how well an alternative hypothesis might explain the same evidence, and the prior
probability of the hypothesis.

The philosophical obstacle that stands in the way of using Bayes's theorem to account for confirmation is the severe difficulty in understanding what a prior probability could be. Consider the prior probability $P(A,B)$. It is the probability that our hypothesis is true regardless of the outcome of our prediction. This is logically prior to the empirical test provided by the hypothetico-deductive method. How are we to make sense of such a probability? One preliminary point is apparent. Prior probabilities fit the description of Hanson's plausibility arguments. Plausibility arguments embody considerations relevant to the evaluation of prior probabilities. They are logically prior to the confirmatory data emerging from the hypothetico-deductive schema, and involve direct consideration of whether the hypothesis is of a type likely to be successful. Plausibility arguments and hypothetico-deductive arguments are essential elements of a logic of scientific inference. We will presently discuss these plausibility arguments with more precision. But for now it suffices that we have located them in the general schema.

Bayes's theorem makes it obvious that we must consider the probability that our prediction would come true even if our hypothesis were false. Other things being equal, the less probable our observational result if the hypothesis is false, the more this observational result confirms the hypothesis.

Popper's view of scientific hypotheses has already been discussed. According to his view the more falsifiable they are, and the more strenuously we have tried to falsify them, the better they are, as long as they survive being falsified. To the extent that hypothetico-deductive theorists have been aware of prior probabilities, they have claimed that hypotheses are better confirmed if they have higher prior probabilities. In other words, if they are plausible. Popper
claims better corroboration for hypotheses that are more audacious and less plausible.

Bayes's theorem says quite unequivocally that plausibility contributes positively to the acceptability of hypotheses. Nevertheless, Popper has a fundamental insight. There is another way in which a hypothesis could run the risk of falsification. Bayes's theorem reveals this. A hypothesis risks falsification by yielding a prediction that is very improbable unless that hypothesis is true. Such a prediction is a daring one, because it is not likely to come out right unless we have hit upon the correct hypothesis. This is reflected in a small value for $P(A,B,C)$. The hypothesis that runs this kind of risk without being falsified gains more in posterior probability than one that runs less of such a risk. This does not mean that the hypothesis itself must be implausible. A small value for $P(A,B,C)$ is perfectly compatible with a large value for $P(A,B)$. There is a good illustration for this type of falsification in the history of optics. Early in the nineteenth century, Poisson deduced from the wave theory of light that the shadow of a disc should have a bright spot in its center. Poisson regarded this as a reductio ad absurdum of the wave theory. But Arago later obtained a positive result when the experiment had been performed. This was an impressive confirmation of the wave theory because the predicted consequence seemed utterly unlikely on any other hypothesis. What was improbable was not the wave theory itself, but the occurrence of the bright spot in the middle of the shadow if the wave theory were not true.

We have seen that Bayes's theorem casts considerable light upon the logic of scientific inference. It provides us with a coherent schema in terms of which we can understand the roles of confirmation, falsification, corroboration, and plausibility. The
theory of scientific inference it yields unifies such irreconcilable views as the hypothetico-deductive theory, Popper's deductivism and Hanson's logic of discovery.

We have so far been concerned with the formal characteristics of Bayes's theorem and what we have said about interpretation has been vague. The formal schema requires prior probabilities. What precisely are prior probabilities?

The Status of Prior Probabilities

The notion of prior probability must be clarified in the light of the interpretation of the probability concept in general. We shall discuss this issue from the standpoint of three leading interpretations.

1. According to the logical interpretation probability is fundamentally an a priori measure of possible states of affairs. The state descriptions provide a list of all possible states of the universe, and weights are assigned to all of them. A scientific hypothesis will be true if certain of these state descriptions hold, and false if others do. The range of the hypothesis is the set of all state descriptions compatible with the hypothesis. The prior probability of the hypothesis is the sum of the values attached to the state descriptions in its range. Accumulated observational evidence enables the calculation of posterior probabilities of hypotheses in accordance with Bayes's theorem. Prior probabilities play an indispensable part in determining the probabilities of hypotheses. But the status of a priori prior probabilities is dubious. So we must reject this interpretation.

2. The personalistic interpretation is based upon a subjective interpretation of probability and makes extensive use of Bayes's theorem. Prior probabilities are totally unproblematic for
the personalist. They are simply degrees of prior belief in the hypotheses before concrete evidence is available. They are subjective plausibility judgments. Evidence may affect this degree of belief, thus issuing in posterior probabilities. Bayes's theorem expresses the relations that must hold among these various degrees of belief. An examination of Bayes's theorem reveals that a prior probability of zero or one determines by itself the same value for the posterior probability. In the remaining cases, prior probability has only a part in determining posterior probability. Under certain assumptions, the role played by the prior probabilities becomes smaller and smaller as observational evidence increases. Personalist theorists give this fact a central place in the arguments. According to them we come to any problem with opinions and preconceptions. Prior convictions of reasonable people can differ considerably. As these individuals accumulate a shared body of observational evidence, differences of opinion will tend to disappear and a consensus of opinion emerge. If the prior opinions do not have the extreme values zero and one their influence will fade in the face of increasing evidence. These individuals need not be genuinely open minded about the various hypotheses. It is enough if their minds are slightly ajar. By showing how the use of Bayes's theorem leads to substantial intersubjective agreement, the personalists argue that the subjectivity of prior probabilities is not pernicious.

This far reaching subjectivism is unacceptable. It is true that satisfaction of the relations established by the probability calculus is a necessary condition for rationality. It is, however, not a sufficient condition. Other requirements for rational belief must be found. All the probabilities that enter into Bayes's theorem in the personalist interpretation are subjective. This
includes the probability \( P(A_B, C) \), that the observational evidence would occur if the hypothesis were true, and the probability \( P(A, \neg B, C) \) that it would occur if the hypothesis were false. All these subjective probabilities may actually be based upon extensive observation and inductive generalization therefrom, but they may also be lacking any foundation whatever in objective fact. Salmon thinks that there is no reason within the personalistic framework to reject as irrational a set of opinions which conflicts with the bulk of experience and dismisses this fact on the grounds that most observation is hallucinatory. Moreover, there seems to be no ground for characterizing as irrational opinions that have arisen out of observation by application of some perverse inductive method. Personalistic theories do not condone misuse of experience in the foregoing ways, but the principles by which they avoid them need to be spelled out, examined, and justified.

3. If one adopts a frequency view of probability, and attempts to deal with the logic of confirmation according to Bayes's theorem (as I do in this essay), then one is committed to regarding prior probability as some sort of frequency - for instance, the frequency with which hypotheses relevantly similar to the one under consideration have enjoyed significant scientific success. The prior probability \( P(A_B) \) is the probability that hypotheses of a certain type are true. The attribute of truth is given directly by the fact that we are looking for true hypotheses. In attempting to choose an appropriate reference class, we are trying to find out what type of hypothesis is likely to be true. A hypothesis that belongs to the class of plausible conjecture is one that has high prior probability. The question is how we are to determine what considerations are relevant to plausibility or prior probability.
Characteristics statistically relevant to the truth or falsity of scientific hypotheses are properties that determine a homogeneous reference class. To evaluate a given hypothesis $H$, we must find a practically or epistemically homogeneous reference class $A$ to which $H$ belongs. $A$ must be a class of hypotheses within which we can say something about the relative frequency of truth. The probability $P(A,B)$ is the probability of truth for hypotheses of this class, and this probability is assigned as a weight to the hypothesis $H$. This prior weight expresses the plausibility of $H$.

Criteria for Plausibility Judgments.

There are three important types of characteristics that may be used as a basis for plausibility judgments. These determine the relevant reference class, but they may also be regarded as criteria of plausibility that hypotheses must confront. Success in meeting a given criterion will classify a hypothesis with other plausible hypotheses; failure will group it with implausible ones.

1. **Formal Criteria.** Scientific hypotheses are proposed against the background of many previously accepted and rejected hypotheses. Newly proposed hypotheses may bear to accepted hypotheses deductive relations that are relevant to their plausibility. If an old hypothesis $H_1$ entails a new hypothesis $H_2$, then the prior probability of $H_2$ is at least as great as the posterior probability of $H_1$. If a new hypothesis $H_3$ is incompatible with an old hypothesis $H_4$, then the prior probability of $H_3$ is no greater than the probability that $H_4$ is false.

2. **Pragmatic Criteria.** There are cases where it is possible to establish a probability relation between the truth of a hypothesis and the circumstances of its discovery. If a religious fanatic who has no training in physics or mathematics were to propose a hypothesis
to replace Einsteinian relativity we would justly place a low estimate on the probability that his hypothesis is true. Such considerations are legitimate only if there is a known probability relation between the character of the individual presenting the hypothesis and the truth of the hypothesis he advances.

3. Material Criteria. Just as relations of entailment or incompatibility can exist between different hypotheses, so too, can there be inductive relations among them. Certain types of hypotheses have been successful; we may legitimately expect new hypotheses that are similar in relevant respects to be successful as well. Analogy with successful hypotheses can serve as the basis for determining the plausibility of hypotheses. The material criteria encompass those respects in which hypotheses may be relevantly similar to one another. We judge the simpler hypothesis more likely to be true because experience has shown us that simpler hypotheses rather than more complex ones have proved to be successful. Or we can distinguish through the kinds of causal processes hypotheses countenance. The physics of Galileo and Newton improved upon the physics of Aristotle by eliminating teleological elements in the latter. The success of non-teleological explanation in physics provided an important precedent for non-teleological evolutionary theories in biology. The success of these theories in biology has provided a strong basis for assigning low prior probabilities to teleological hypotheses in psychology and sociology.

These examples of material criteria provide some idea of the kind of plausibility argument falling under that head. We have seen that three criteria provide the grounds on which we can legitimately decide what kind of hypotheses are likely to succeed. It is experience that helps determine prior probabilities.
It may appear that the whole discussion has done little to show how precise values can be assigned to prior probabilities. Numerical precision is not required. Bayes's theorem will be applicable if we can merely judge whether the hypothesis is totally improbable. If the prior probability can be taken as zero the hypothesis can be disqualified from further consideration. A zero value for \( P(A,B) \) settles the question because, in that case, \( P(A,C,B) \) is also zero. Even a very small prior probability of the hypothesis leaves the question of its posterior probability open.

The frequentist - and that is the position we adopt - is committed to regarding prior probability as some sort of frequency. So the question of the plausibility of hypotheses has something to do with our experience in dealing with hypotheses of similar types. Thus the reason we would place a low plausibility value on teleological hypotheses is related to our experience in the transitions from teleological to mechanical explanations in the physical, biological, and to some extent, the social sciences. To turn back towards teleological hypotheses would be to go against a great deal of scientific experience about what kinds of hypotheses work well scientifically.
THE INTELLIGENT DESIGN HYPOTHESIS

The aim of the design argument is to show that the universe which exhibits a high degree of orderliness is very probably the result of intelligent design.

In attempting to assign a probability to the hypothesis that the universe was created by an intelligent being, we are dealing with a unique event. How are we to assess the probability of a single event? According to the frequency interpretation, probability is a relation between two classes. This fact is reflected in the notation, \( P(A, B) \) where expressions for two classes have been incorporated. The class mentioned first is the reference class; the other is the attribute class. The problem of the single case is the problem of selecting the appropriate reference class.

An insurance company, in determining what premium an individual should be charged for his automobile insurance, assigns him to a category of drivers who are similar to him in relevant aspects. Reichenbach would have us choose the narrowest reference class for which reliable statistics are available. Salmon says, instead, that the single case should be referred to the broadest homogeneous reference class of which it is a member. In either formulation the intent is fairly straightforward. Probability has to be established inductively. So we must have enough instances to be able to make inductive generalizations. We do not want to refer our single cases to classes that are too narrow because if we do, we will not have enough evidence on which to base our inference. At the same time, we want our reference classes to
contain other relevant cases and not irrelevant ones. The key concept is statistical relevance. Suppose we ask for the probability that a given individual $x$ has a characteristic $B$. We know that $x$ belongs to a reference class $A$ in which the limit of the relative frequency of $B$ is $p$. If we can find a property $C$ in terms of which the reference class $A$ can be split into two parts $A\cdot\neg C$, such that $P(A\cdot C, B) \neq P(A, B)$ then $C$ is statistically relevant to the occurrence of $B$ within $A$. $C$ must be the sort of property whose occurrence in an individual can be detected without knowing whether that particular entity also has the property $B$. If there is no such property $C$ by means of which to effect a relevant subdivision of $A$ with respect to the occurrence of $B$, $A$ is said to be homogeneous with respect to $B$. Let us apply this consideration to the design argument.

Let us now attempt an assessment of the probabilities required by Bayes's theorem. We will begin with the prior probabilities $P(A, B)$ and $P(A, \neg B)$, the probability that a case of coming-into-being is an instance of the operation of intelligence, and the probability that a case of coming-into-being is an instance of the operation of something other than intelligence.

We cannot perform an induction by enumeration on observed births of universes in order to draw a direct conclusion about the creation of our own. As Peirce has remarked, universes are not as plentiful as blackberries. "Have worlds ever been formed under your eye, and have you had leisure to observe the whole progress of the phenomenon, from the first appearance of order to its final consummation?" asks Philo in Hume's Dialogues Concerning Natural Religion. The answer is negative. We can only make an indirect inference from the origins of other types of entities whose beginnings we can observe. So let us proceed on that basis.
The first step in applying Bayes's theorem is finding meanings for the terms that occur in it. Let 'A' denote the class of instances of entities coming into being. We will consider a broad class so as not to bias our evidence. The class A will include, for example,

(i) The formation of a fetus when a sperm and egg unite.
(ii) The building of a house that was designed by an architect.
(iii) The formation of ice as water freezes.
(iv) The growth of a tree from a seed.
(v) The carving of a gully by flowing water.

Let 'B' designate the class of instances in which intelligence operates.

(i) Proving a mathematical theorem.
(ii) Designing and making a watch.
(iii) Composing a piece of music.
(iv) Designing a house.

'C' will be taken to represent the class of entities which exhibit order. For example,

(i) A watch.
(ii) The solar system.
(iii) A living organism.

All the foregoing examples are entities which obviously exhibit some sort of order. We shall return to the concept of order below, to analyze it with greater precision.

Now that we have assigned meanings to the letters in Bayes's theorem we are in a position to interpret all of the probability expressions that appear in it.

\[
P(A,B) = \text{the probability that a case of coming-into-being is an instance of the operation of intelligence.}
\]

\[
P(A,\overline{B}) = \text{the probability that a case of coming-into-being is}
\]
an instance of the operation of something other than intelligence.

\[ P(A \cdot B \cdot C) = \text{the probability that something produced by intelligence exhibits order.} \]

\[ P(A \cdot \overline{B} \cdot C) = \text{the probability that something produced by some agency other than intelligence exhibits order.} \]

\[ P(A \cdot C \cdot B) = \text{the probability that something which comes into being and exhibits order was produced by intelligence.} \]

This last probability is the one we seek. Proponents of the design argument maintain that it is very high.

When we consider living organisms, both animal and vegetable, we see that biological reasons account for their generation in a very large number of instances. Birds build nests, bees make honeycombs, spiders spin webs and a whole host of other animals and insects build places to live in, store their food, and give birth to their young. The coming into being of such things as bird nests, spider webs, and honeycombs can be accounted for by the principle of instinct. The principle of mechanical causation accounts for the formation of snowflakes, and crystals, molecules and atoms, galaxies and solar systems. This principle obviously operates with great frequency upon the earth and, for all we know, everywhere else in this vast universe. Our universe contain perhaps 10 billion galaxies. Each of these contain, on the average from 10 to 100 billion stars. Who can number the atoms formed in the interiors of these stars? It is estimated that our earth alone contain about \(10^{50}\) atoms. There are, it must be acknowledged, many unanswered questions in cosmology and astrophysics concerning the formation of galaxies, stars, and atoms. But we have achieved some scientific understanding of these processes. They appear to be mechanical. There is, consequently,
a great deal of evidence in support of the conclusion that the number of instances in which mechanical causation works is overwhelmingly greater than all of the rest combined; and the rest includes all these types of human artifacts that arise from intelligent design. Our evidence indicates that \( P(A,B) \) is low, and that \( P(A, \overline{B}) \) is high. We now have a rough assessment of the prior probabilities.

We must give a closer specification of the particular case with which we are dealing. We need to take into account the type of order the universe exhibits. But first we must examine the nature of the intelligent creator hypothesized by the proponent of the design argument.

If we pay heed to experience it is impossible to assign a high probability to the hypothesis that the world, if created as a divine artifact, was the product of a God bearing any resemblance to the theists' conception. Experience suggests that a universe of such sheer magnitude was not created by a unitary being. Where human artifacts are concerned, the larger the project, the more likely it was to have been executed by a group of people. Moreover, even very complex machines are sometimes made by dull artisans who merely copy the work of others. As Humes put it in the mouth of Philo, if we take into consideration the magnitude of the machine, and the imperfections in its construction, for all we know the world may have been created by a juvenile deity who had not mastered his trade, or a stupid deity who could only make bad copies, or a committee of deities or a senile one who had lost the knack by the time he got round to making the world. All of our experience suggests that the world was not created by a deity bearing any resemblance to the hypothesized deity of the design argument.

Worse still, the God of traditional theism is regarded as pure spirit. In no instance within our experience has a disembodied
intellect produced any artifact, whether it might have exhibited order or not. To the best of our knowledge disembodied intelligence has never operated in any fashion. We are forced to conclude that for such an intelligence, \( P(A,B) = 0 \) and that \( P(A,B,C) \) is undefined.

Theists believe that the universe gives evidence of a creator who is intelligent powerful and benevolent. The addition of the moral attributes has an adverse effect upon the likelihood of the theistic hypothesis. Other things being equal, lowering the likelihood results in a reduced posterior probability. The likelihood \( P(A,B,C) \) is very low because the world seems utterly indifferent to values. Rain falls upon the just and the unjust. Evil abounds. Mankind is plagued by misery and suffering. Nuclear, chemical and biological warfare would be quite indiscriminating in their destruction. The problem for the design argument is not that of reconciling the existence of an omniscient, omnipotent and perfectly benevolent God with the existence of evil on earth. Rather the proponent of natural religion must maintain that the world as we know it is positive evidence for the claim that it was created by a wise, powerful, and benevolent God. A scientific hypothesis is tested by deducing observational consequences that would hold if the hypothesis were true, and then observing to see whether such consequences obtain. In the eleventh dialogue Philo asks:

"Is the world, considered in general and as it appears to us in this life, different from what a man ... would, beforehand, expect from a very powerful, wise, and benevolent Deity? It must be strange prejudice to assert the contrary. And from thence I conclude that, however consistent the world may be, allowing certain suppositions and conjectures, with the idea of such a Deity, it can never afford us an inference concerning his existence. The consistency is not absolutely denied, only the inference" Hume, Dialogues Concerning Natural Religion, p.96.

Philo sees four rather obvious ways in which an all powerful, wise and benevolent God could have reduced the amount of evil in the
world if he had wanted to (Ibid., pp. 96-102). (i) Man need not have been endowed with the capacity for pain. (ii) God need not govern the world by inviolable laws. A little tinkering now and then could be quite beneficial: an extra large wave upon a warship out to inflict harm on helpless people: a little calming of the waters for a ship on an errand of mercy. (iii) Mankind and the other species have been so frugally endowed with capacities as to make their existence hazardous and grim. Why could not the creator have endowed them more generously, for example, with such capacities as industry, energy, and health? (iv) A further source of "misery and ill of the universe is the inaccurate workmanship of all the springs and principles of the great machine of nature ... they are, all of them, apt, on every occasion, to run into one extreme or the other. One could imagine that this grand production had not received the last hand of the maker so little finished is every part, and so coarse are the strokes with which it is executed" (Ibid., p. 101).

What Philo is asking is whether this world is what we would antecedently expect of an omnipotent, omniscient, and benevolent God.

"Did I show you a house or palace where there was not one apartment convenient or agreeable; where the windows, doors, fires, passages, stairs, and the whole economy of the building were the source of noise, confusion, fatigue, darkness, and the extremes of heat and cold, you would certainly blame the contrivance, without any further examination. The architect would in vain display his subtility, and prove to you that, if this door or that window were altered, greater ills would ensue. What he says may be strictly true: The alteration of one particular, while the other parts of the building remain, may only augment the inconveniences. But still you would assert in general that, if the architect had had skill and good intentions, he might have formed such a plan of the whole, and might have adjusted the parts in such a manner as would have remedied all or most of these inconveniences. His ignorance, or even your own ignorance of such a plan, will never convince you of the impossibility of it. If you find any inconveniences and deformities in the building, you will always, without entering any detail, condemn the architect" (Ibid., pp. 95-96).

Given the hypothesis that God is supremely wise, good and powerful
there is a very low probability, $P(A,B,C)$, that he would create a world such as this. This is not to say that the world does not exhibit order but that the kind of order it exhibits is not what would be expected as a result of creation by a wise, powerful, and benevolent creator.

There is a logical reason why the proponent of the design argument cannot abandon consideration of the moral attributes. When one asks what sort of order the world exhibits which gives such strong evidence of intelligent design, the answer one receives is "the adjustment of means to end." Such considerations cannot be brought to bear unless we have some conception of what ends the means are designed to serve. If the ends are justice, mercy, and benevolence, one sort of order would count for intelligent design. If the end is to secure from mankind blind unreasoning adulation of and obedience to a Deity who says in one breath, "Thou shalt not kill," and in another, "Slay and spare not," another kind of order would be suitable. If the creator is mainly concerned with winning bets with Satan, as in the book of Job, another kind of order would be appropriate. Thus, the moral attributes are inextricably bound up in the argument directed towards his natural attributes. As Hume has said, "Reason is, and might only to be, a slave of the passions, and can pretend to no other office but to serve and obey them". (A Treatise of Human Nature, Bk.ii. Part iii. Sect.iii).

The theist believes that the order he finds in the world furnishes him with compelling evidence of intelligent design. The consideration that the order exhibited by living organisms comes from biological generation does not seem to the proponent of the design argument to undermine his position. Indeed such biological wonders reinforce his argument. Paley addresses this issue:
"Suppose, in the next place, that the person, who found the watch, should, after some time, discover, that, in addition to all the properties which he had hitherto observed in it, it possessed the unexpected property of producing, in the course of its movement, another watch like itself; (the thing is conceivable;) that it contained within it a mechanism, a system of parts, a mould for instance, or a complex adjustment of lathes, files and other tools, evidently and separately calculated for this purpose; let us inquire, what effect ought such a discovery to have upon his former conclusion?

The first effect would be to increase his admiration of the contrivance, and his conviction of the consummhel skill of the contriver .... If that construction without this property, or which is the same thing, before this property had been notice, proved intention and art to have been employed about; still more strong would the proof appear, when he came to the knowledge of this further property, the crown and perfection of the rest (William Paley, Natural Theology, pp. 8,9).

"The conclusion which the first examination of the watch, of its works, construction, and movement suggested, was, that it must have had, for the cause and author of that construction, an artificer, who understood its mechanism, and designed its use. The conclusion is invincible. A second examination presents us with a new discovery. The watch is found in the course of its movement, to produce another watch, similar to itself; and not only so, but we perceive in it a system of organization, separately calculated for that purpose. What effect would this discovery have, or ought it to have, upon our former inference? What, as hath already been said, but to increase, beyond measure, our admiration of the skill, which had been employed in the formation of such a machine? Or shall it, instead of this, all at once turn us round to an opposite conclusion, viz. that no art or skill whatever has been concerned in the business, although all other evidences of art and skill remain as they were, and this last supreme piece of art be now added to the rest? Can this be maintained without absurdity? Yet this is atheism (Ibid. pp.12,13)

In these passages, Paley deals with the most fundamental issue concerning natural religion and the design argument. For the theist, the order which arises out of biological generation is only further evidence of intelligent design. But, within our experience, all instances of intelligent design issue from biological organisms - biological generation always lies behind intelligence. We do not have experience of intelligent creation as a prior source of biological generation. Why then does the defender of the design argument insists contrary to experience, that intelligent design lies behind the operation of any other principle which generates order. The
answer lies in a teleological conception of order. Cleanthes sees the universe as a "great machine", composed of a prolific array of "lesser machines", all of which are characterized by "the curious adapting of means to end". This is a theme conspicuous in Paley who believed that the parts work together to achieve some useful end. Defenders of this argument seem to equate order with design. But this procedure begs the question. It appeals to an a priori principle according to which order is inseparably attached to thought and that it can never of itself or from original unknown principles belong to matter. The success of classical mechanics, and Darwinian evolutionary theory, banished teleological principles from nature. Galileo and Newton were responsible for removing the Aristotelian teleological conception from physics. The beautiful order of the solar system could be reduced to mechanical principles. The net result of the scientific revolution was to present a picture of the world which exhibits a wonderful order and simplicity. The world is seen as a collection of material particles which respond to forces in accordance with Newton's three simple laws of motion. The further development of classical mechanics, which reached its zenith at the end of the nineteenth century, revealed that other forces, particularly electric and magnetic forces, also conform to simple and precisely specifiable laws. At the turn of the century, it appeared that all natural phenomena could be explained in terms of these fundamental principles of classical physics. Darwinian evolutionary theory purged biology of teleological principles. Darwin showed how nonpurposive factors of chance mutations and natural selection could lead to the evolution of the species, and in the process of evolution these species which are best adapted to compete for food and reproductive opportunities would evolve. It is true that evolutionary biology does pose many
problems, but there is no reason to think that a retreat into teleology is the answer. Twentieth century molecular biology has added to our understanding of the mechanisms of heredity by showing precisely how order is reproduced.

The constituents of the world behave in accordance with simple laws. This is a kind of order, but it is not the only kind of order nature exhibits. A random and totally disorganized system of material particles would obey the same laws as a finely constructed machine. Heaps of metal in junk yards obey the same physical laws as the most complex machines. The universe is not a cosmic heap of unorganized parts. Eighteenth and nineteenth century writers could not characterize this more satisfactorily than in terms of the adjustment of means to ends, and the similarity of the universe to man-made machines. We cannot deal with the means-end relationship if we do not have some independent evidence concerning the nature of the end which is supposed to be served. The similarity of the world to a machine can easily be challenged. Philo does precisely this when he says that the world resembles more a plant or a vegetable than a watch or knitting loom.

Scientific developments since about the middle of the last century have significantly clarified the concept of order. In dealing with physical problems closely related to practical concerns about the efficiency of machines, physicists and engineers created the science of thermodynamics. They established a viable concept of energy, discovered laws relating the various forms of energy to each other, and developed the concept of entropy. Entropy is essentially a measure of the unavailability of energy to do mechanical work. Given two physical systems, each with the same amount of energy, the one with the lower entropy is the one from which it is possible in principle to extract the greater amount of work. All machines in the course
of their operation, tend to dissipate some of their energy in useless forms. They are subject to the universal law of degradation of energy, the second law of thermodynamics. The discovery of this fundamental law of nature that entropy tends to increase has led to considerable speculation about the "Heat death" of the universe.

Towards the end of the nineteenth century, thermodynamics was given a theoretical foundation in statistical mechanics, and the concept of entropy was given a statistical interpretation. Low entropy was shown to be associated with non-random, highly ordered arrangements, which are relatively improbable. High entropy is associated with random, unordered arrangements which are relatively probable. If, for example, we have a container with 80 molecules in it - 40 oxygen molecules and 40 nitrogen molecules - the arrangement would be highly ordered and non-random if all of the oxygen molecules happened to be on the left-half of the container and all of the nitrogen molecules happened to be in the right half. The probability of the molecules sorting themselves out in this fashion in the course of their random motions is about $10^{-24}$. Other arrangements where the molecules are dispersed in a more disordered manner and at random are vastly more probable.

The problem of the design argument can be restated in terms of thermodynamic considerations. If we may speak in terms of the entropy of the whole universe, it would seem that the order exhibited by the universe can be described by saying that the entropy of the universe is now relatively low, and that it has been even lower in the past. To say that the entropy is low is tantamount to saying that the universe contains large stores of available energy.

The question of creation can now be posed again. The universe is a physical system in a relatively low entropy state. Is there
strong reason to claim, by virtue of this fact, that it must have been produced by intelligent design?

According to the modern statistical interpretation of the second law of thermodynamics, the entropy in a closed physical system is very probably high. It is not impossible for the entropy of such a system to decrease as a result of a mere statistical fluctuation. But such instances are very improbable. It is also possible for a closed physical system to be in a state of low entropy as a result of a recent interaction with its environment. For example, a thermos bottle containing tepid water with ice cubes floating in it is in a low entropy state. In course of time the ice cubes will melt, the water will cool, and the whole system inside the insulated container will arrive at a uniform temperature. This is a state of higher entropy, upon finding a thermos whose contents were in the lower of these two states, we would infer without hesitation that it had recently been put into that state by an outside agency - in this case, by a person who removed the ice cubes from a refrigerator and deliberately placed them in the thermos with water. This is a case of low entropy resulting from interaction with an intelligent planner who put the ice cubes into water to fulfill a conscious purpose. Low entropy states which are the result of interaction with the environment do not always involve human intervention. A hailstorm on a summer day may deposit pieces of ice in the luke warm water of a swimming pool.

It is more desirable to reformulate the problem of intelligent design in terms of the concept of entropy. The vague concept of order can thus be avoided and there will not be as strong a temptation to beg the question by identifying low entropy with conscious design as there was to identify order with purpose and "the adjustment of means to ends". We may look around the world, surveying the physical
systems which come into being with low-entropy states, to ascertain what percentage of them are created with conscious design. The result will be similar to those already discussed. An exceedingly small proportion of low-entropy systems results from an interaction with the environment which involves any conscious purpose or design. We are forced to conclude that the probability, \( P(A, B, C) \) is much larger than the probability \( P(A, B, C) \). In other words it is a lot more likely that some cause or causes other than an intelligent creator explains the order we find in the universe.

Today we are in a position to claim some physical knowledge about the evolution of the universe. There is reason to believe that, sometime between ten and twenty billion years ago, the universe consisted of a compact concentration of energy which exploded with incredible violence. This 'big bang' theory is supported by observed red-shifts of light from distant galaxies and by the more recently discovered cosmic microwave background radiation. It would seem likely that statistical fluctuations in the rapidly expanding fireball gave rise to stable inhomogeneities, by gravitational attraction of neighbouring matter, led to the formation of galaxies. Further concentrations within the galaxies led to the formation of stars. Our universe contains about ten billion such galaxies. Each of these is an inhomogeneous concentration of energy. Each galaxy contains about ten billion stars. These again are inhomogeneous concentrations of energy. We can cite 100 billion billion systems which came into being in states of low entropy. Where can we find like numbers of systems created in low entropy states by conscious human intervention? We have some knowledge of how atoms are formed in the interiors of stars. The earth alone contains about \( 10^{50} \) atoms, each a highly organized low-entropy system, and each, to the best of our knowledge, brought into
being without conscious intent. The low entropy systems created without intelligent purpose outnumber overwhelmingly those which involve intelligent design. We can conclude that the design argument is very improbable.

CONCLUSION

Let us now take stock of the whole situation. We have made plausibility assessments of the probabilities required for the application of Bayes's theorem.

Let us consider the prior probability $P(A,B)$ — i.e. the probability that the universe was created by an intelligent being. Since we are dealing with a unique event we must refer it to a broad reference class and so can only make an indirect inference from the origins of entities whose beginnings we can observe. All of our experience suggests that $P(A,B)$ is overwhelmingly greater than $P(A,B)$. The prior probability of the intelligent design hypothesis is very low.

We can also say quite confidently that the probability $P(A,C,B)$ — that an unspecified entity, which came into being and exhibited order, was produced by intelligent design — is quite low. This conclusion can be reinforced even more directly when we consider all the entities with which we are acquainted which come into being exhibiting order. If we count the relative frequency with which such entities were the result of intelligent design, we see that it is rather low. This conclusion holds even when we exclude such items as galaxies and atoms on the grounds that we are not very sure how they are created. There is still a vast numerical preponderance of such occurrences as animal reproduction, growth from seeds, formation of crystals, and spinning of spider webs over the relatively few instances in
which watches, houses, and ships are built by man.

When we consider the nature of the creator hypothesized by proponents of the design argument it becomes clear that the kind of order the world exhibits is not what would be expected as a result of creation by a wise, powerful and benevolent deity. If the intelligent design hypothesis is true it makes the facts to be explained quite improbable.

We have to assign a low plausibility value to the intelligent design hypothesis. Experience teaches us that mechanical hypotheses work better than teleological ones. The history of science reveals a transition from teleological to mechanical explanations in the physical, biological, and even the social sciences. Furthermore the teleological intelligent design hypothesis fits very badly with recent scientific developments.

In brief:

(i) The intelligent design hypothesis is antecedently implausible.

(ii) It makes the facts to be explained quite improbable if it is true.

(iii) There are plausible alternative hypotheses which make the facts to be explained highly probable.

Under these circumstances, only gross prejudice would make one retain the intelligent design hypothesis.
BIBLIOGRAPHY


