MODELS OF SCIENTIFIC EXPLANATION

A Thesis

by

PETER ANDREW SUTTON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

May 2005

Major Subject: Philosophy

MODELS OF SCIENTIFIC EXPLANATION

A Thesis

by

PETER ANDREW SUTTON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Approved as to style and content by:

Robert Burch (Chair of Committee)	Colin Allen (Member)
Mark Everett (Member)	Robin Smith (Head of Department)

May 2005

Major Subject: Philosophy

ABSTRACT

Models of Scientific Explanation. (May 2005)

Peter Andrew Sutton, B.A., Southwest Texas State University

Chair of Advisory Committee: Dr. Robert Burch

Ever since Hempel and Oppenheim's development of the Deductive Nomological model of scientific explanation in 1948, a great deal of philosophical energy has been dedicated to constructing a viable model of explanation that concurs both with our intuitions and with the general project of science. Here I critically examine the developments in this field of study over the last half century, and conclude that Humphreys' aleatory model is superior to its competitors. There are, however, some problems with Humphreys' account of the relative quality of an explanation, so in the end I develop and defend a modified version of the aleatory account.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

An explanation is simply an answer to a "Why?" question. "Why does the sun appear to rise and set?" "Why did you step on my foot?" "Why do bad things happen to good people?" "Why did that red warning light come on?" "Why do objects fall to the ground when released?" An answer to any of these questions would be an explanation of the fact or event in question. As the questions indicate, the sorts of facts that need explaining can vary greatly, from the minor to the momentous, from the broad to the specific, and from the trivial to the significant. With such a diversity of things to be explained, an analysis of the notion of "explanation" would be a vast undertaking indeed. I do not plan to offer any such analysis.

What I will do, in the present paper, is study the various ways that "explanation" is understood within the physical sciences. By "physical sciences," I have in mind primarily physics (in all its many forms), but also all those fields that understand the world in purely physical terms. These include chemistry, geology, astronomy, and meteorology. Biology and medicine, along with such sub-disciplines as genetics and neurology, are to be included as physical sciences, but only to the extent that their explanations do not involve teleology. For instance, a proper biological explanation (in the sense used in this paper) of a bird's consumption of a particular type of food might be

---

This thesis follows the style of *The Chicago Manual of Style.*

that it was genetically predisposed to eat such things, or that it was conditioned by such-and-such a procedure to eat such things. However, it would not be appropriate, for the purposes of this paper, to say that the bird ate the food because he "likes the taste," or even "in order to survive." I make no claims about the truth, falsity, or importance of such teleological explanations; I simply do not include them amongst what I shall call *scientific* explanations. So, with these boundaries in place, I think we can reasonably expect to be able to develop a useful and consistent idea of *scientific* explanation, even if our notion of explanation *in general* is too broad to approach.

Having said which, there are a number of competing theories of what counts as scientific explanation, and there is by no means a consensus within the philosophical or scientific communities about which theory is most appropriate. I will argue, in what follows, that Paul Humphreys' theory of aleatory explanation is superior to the other theories within the current literature on the subject, but that an important improvement needs to be made to it. Specifically, the way that Humphreys handles the distinction between "good" and "bad" explanations has counterexamples, and must be changed in order to preserve the intuitive correctness of the theory.

Before I start the paper proper, I want to present some of my working assumptions and define a few of the terms that will be important.

First, I am assuming, for the purposes of this paper, that our current best scientific theories are, for the most part (and as far as they go), correct. I am of course aware that scientific theories are not set in stone or immune to error, but if we are to take science seriously as a discipline (which I am assuming we should) then we should not allow our philosophical theories to contradict it in any serious way. So whatever theory of scientific explanation is most satisfactory had better accord reasonably well (though perhaps not perfectly - it is fine to hold science to be in error *sometimes*) with our current best scientific theories.[1]

Second, I am assuming that explanations can only be offered for events that have in fact occurred. In other words, I assume that the only "why?" questions that may reasonably be asked are those where the statement that needs to be explained is true. So "why P?" only makes sense when P is true. If there are any exceptions to this rule, then this thesis makes no claim to deal with them.

Third, I am assuming that there is in fact a *best* theory of scientific explanation, and that it applies equally well to each branch of physical science. I present a number of different explanatory models in this paper, and some might argue that as each has different strengths and weaknesses, that each should be eclectically applied according to where it best *fits*, and that there is no *one* correct way of explaining. I will not argue explicitly against this view, but I am assuming its falsity for the purposes of my arguments.

Fourth, I will be discussing explanation within the physical sciences, but this does not mean that our common sense understanding of explanations has no place in my arguments. I will be assuming that the word "explanation" in science does not differ too

[1] This is not to say that our theory of explanation should affirm the *truth* of any particular scientific theory, only that it should be useful when applied to such theories. Our theory of explanation should really be as neutral as possible with regard to the *truth* of any particular scientific claim.

greatly in meaning from our ordinary use of the word, even if there are certain minor alterations. If this assumption were incorrect, then scientific explanations would not be as interesting and useful to us as they seem to be. This assumption is important because it allows for the use of philosophical thought-experiments in my argumentation.

As far as my terminology goes, the term "explanandum" refers to the sentence describing the event to be explained. The term "explanans" refers to a sentence that explains (in whole or in part) the explanandum. I will take "explanation" and "scientific explanation" to be equivalent terms, each referring to "explanation within the physical sciences[2]," and all my arguments shall be restricted to explanation thus understood.

[2] Which are, in turn, defined above.

CHAPTER II

THE STANDARD MODELS OF EXPLANATION

In the introduction I said that explanations are usually taken to be answers to "why?" questions. This is a good starting point, but tells us little about what such answers must look like, and what makes them good answers. In this chapter I will present several of the standard models of scientific explanation. I cannot hope to be comprehensive, but I think the five models I have chosen are rather indicative of the different types that are in the literature.[3] First, I will give the basic idea behind each model and present its formal apparatus (when applicable). I will then show the particular strengths of the model in question and how it supports our intuitions or supports the practice of science.

In the end, however, each of these five models has some rather serious inadequacies. Once all the models have been presented in full, I will discuss the difficulties that each presents us. By properly understanding where each of these models goes wrong, we can hopefully gain a better understanding of what conditions an adequate model of scientific explanation must meet. In the following chapter, I argue that Paul Humphreys' theory of aleatory explanation meets all these conditions and is therefore superior to the alternatives in the literature.

[3] For instance, Coffa's dispositional theory of explanation is much like Hempel's I-S explanation with an "all other things being equal clause" in the probabilistic law, and Bromberger's account of explanation is much like Lewis', but with a less explicit reference to causality. The only model I am aware of that is different *in type* from those presented here is Bas van Fraassen's pragmatic account, which I go into in detail in my chapter on contrastive explanation and the symmetry thesis.

The Deductive Nomological Model

Carl G. Hempel and Robert Oppenheim, in their 1948 essay "Studies in the Logic of Explanation," developed the Deductive Nomological (D-N) model of scientific explanation[4], which has been the most influential model of explanation in the last fifty years. The D-N model takes a scientific explanation to be a deductive argument in which the explanandum is inferred from some general laws of nature and some particular facts relating to the case in question. The explanans must be empirically verifiable and the laws in question must be necessary in order to reach the conclusion. So a scientific explanation has the following form:

(D-N)* $L_1, L_2, \ldots, L_r$

$$\frac{C_1, C_2, \ldots, C_k}{E}$$

*Where $L_1, L_2, \ldots, L_r$ are general laws, and $C_1, C_2, \ldots, C_k$ are particular facts about the case in question. E is the explanandum event[5]

This model can be expressed more simply (but less generally) in the following manner:

(D-N) $\forall x(Fx \supset Gx)$

$$\frac{Fi}{Gi}$$

It might initially be objected that since the simplest deductive argument form of all is 'P $\vdash$ P', and since we already know in scientific explanations that the explanandum is true, any event can be given a D-N explanation that is trivially circular. This would be

[4] Carl Hempel and Robert Oppenheim. "Studies in the Logic of Explanation."
[5] Carl Hempel "Deductive-Nomological Versus Statistical Explanation." 87-90.

a mistake, however, since it was stipulated that a D-N explanation must have a true law of nature in the premises that is *necessary* for the inference to go through. This implies that the explanandum cannot be included amongst the premises, so this form of circularity is avoided.

The deductive form of a D-N explanation, along with the requirement that the premises be true, means that any such account will give us a sound proof that the conclusion (the explanandum) is true. Our best theories (provided that the laws involved in them are true), along with the facts about our experiments, are the premises from which one can deduce the occurrence of the explanandum. The fact that it does so by the use of some law of nature is what makes this a *scientific* explanation. Hempel and Oppenheim consider an explanation of this sort to be an adequate answer to any "why?" question asked of a deterministic physical system.

The Inductive-Statistical Model

Hempel realized that D-N explanations would only take us so far. There are aspects of the world that we really can only understand statistically, like medical explanations (of why certain people became infected with a given virus) and biological explanations (of why certain species become the dominant ones in a certain ecosystem). Furthermore, there are aspects of the world, like the phenomena of quantum physics, which our best science says *just are* statistical, and are not merely *understood* statistically. And so Hempel developed what he calls the inductive-statistical (I-S) model to augment the explanatory power given us by the D-N model. Here is the form of the I-S model:

(I-S)* p(Gx | Fx) = r
Fi
======== [r]
Gi

This is an adequate explanation relative to a given "knowledge situation" K only if r is close to 1 and both premises are contained in K.[6]

The above model reads basically like this. As far as we know, the probability that an F will be a G is r (which is close to 1), and this i is a case of an F. Therefore, i is likely to be (has probability r of being) a G. In I-S explanations, the conclusion is not deductively implied by the premises, but is rendered *highly expectable* (given our knowledge situation) to degree r. We cannot show that the explanandum was certain to occur, but we can show that it was very likely, given what we know, which Hempel regards as an adequate explanation. However, Hempel realized that we are not quite done, and that a further requirement is necessary to make this model acceptable. Consider the following example. John has been infected with malaria plasmodium. Generally speaking, the probability of contracting malaria in such a situation is, say, .9, and John does in fact contract malaria. Now consider the following argument:

(A) p(John contracts malaria | John is infected with malaria plasmodium) = .9
John is infected with malaria plasmodium
=== [.9]
John contracts malaria

[6] Carl Hempel. "Maximal Specificity and Lawlikeness in Probabilistic Explanation." 146-50

Or, in symbols:

p(Mx | Ix) = .9
Ij
===========[.9]
Mj

This is an I-S explanation, as the premises are both contained in K (our knowledge-situation, which contains relevant statistical information about malaria, in addition to some non-statistical information about John's having been infected), and the premises confer a high degree of expectability on the conclusion. As John has in fact contracted malaria, it would seem that Hempel would regard this as an adequate explanation of the fact.

But imagine we alter the case somewhat. Say our knowledge-situation includes information to the effect that John has acquired the sickle hemoglobin gene from one, but not both, of his parents. This reduces John's chances of contracting malaria considerably, such that the probability he will contract the disease after having been infected is, say, .05. Therefore, the probability of his *not* getting the disease in the case described above is .95. So let Hx stand for 'x is a heterozygote in regard to the sickle hemoglobin gene[7]'. Now imagine that John did *not* contracted malaria. The following argument might be constructed:

(B) p(¬Mx | Ix & Hx) = .95
Ij & Hj
=================[.95]
¬Mj

[7] That is, 'x has acquired the sickle hemoglobin gene from one, but not both, parents'.

(B) is an apparently adequate I-S explanation of John's *not* getting malaria. (A) is an adequate I-S explanation of John's *getting* malaria. We could imagine all the premises for both arguments being contained within a single knowledge-situation, and both confer a high probability upon their respective (mutually exclusive) conclusions. So the exact same body of knowledge can explain John's getting malaria if he does and explain his not getting malaria if he doesn't. Hempel sees this as a huge problem. If an event E is probable and explicable in a given knowledge-situation, then it is (according to Hempel) unthinkable that ¬E also be probable and explicable in that same situation[8]. However, it would seem that statistical explanations are full of such peculiarities. Most objects in our universe can be accurately categorized in many different ways. Within almost any class there are many sub-classes, each of which might have different probabilities with regard to a given property or event than the broader class. John is a member of the class of all people who have been infected with malaria plasmodium, but is also a member of a subclass that consists of those people who have been infected but are heterozygotes in regard to the sickle hemoglobin gene. The probability of contracting malaria in the broad class is much higher than in the subclass, and so, depending on which facts we look at, we might be lead to believe that he is very likely or very unlikely to contract malaria.

What Hempel does to resolve this is to restrict what sorts of classes count as statistically relevant in our I-S explanations. He puts a restriction on the I-S model that

[8] I agree with Hempel (as I think anybody with a reasonable understanding of probability would) that E and ¬E could not both be *probable*. It is with his claim that they could not both be *explicable* that I disagree. See, in particular, my chapter on contrastive and non-contrastive explanations.

will *make* us pay attention to *both* John's genetic history *and* the general facts about humans' infection with and contraction of malaria. What he introduces is called the *Requirement of Maximal Specificity*. The model, with its new condition, looks like this:

(I-S) p(Gx | Fx) = r
Fi
==========[r]
Gi

Where r is close to 1 and both premises are contained in K, this constitutes a probabilistic explanation relative to K only if:

(RMS) For any class F_1 for which K contains statements to the effect that F_1 is a subclass of F and that F_1i, K also contains a probabilistic-statistical law to the effect that $p(Gx \mid F_1x) = r_i$, where r_i=r unless the law is a theorem of probability theory.

What the requirement of maximal specificity says, simply put, is that if the thing in question (i) is a member of any subclasses of the class we are subsuming it under (F), then those subclasses had better not be statistically relevant to the property in question (G). So (A) above is not an acceptable I-S explanation since John *is* a member of a subclass of I (those infected), namely the class {H & I} (those infected who are heterozygotes in regard to the sickle hemoglobin gene), and the probability of contracting malaria for the member of that subclass is different that it is for the broader class. (B) *would* count as an I-S explanation, but only if John is not, according to K, a member of any further subclasses of I (or of {H & I}) in which the probability of contacting malaria differs from that of the class {H & I}.

The additional bit about the law in question not being a theorem of probability calculus is meant to address that problem that arises when we consider that we only ever attempt to explain things that are, in fact, the case[9]. Because of this, all objects about which we offer explanations are members of subclasses of things that have the explanandum-property. This, we shall see, poses a temporary problem for Hempel. If John does not have malaria, we might use (B) as an explanation of this fact. But unfortunately, John is a member of a subclass of {H & I} in which the probability of contracting malaria differs greatly from that of {H & I}, namely the subclass of all those who do not contract malaria: {H & I & ¬M}. In the class {H & I & ¬M}, the probability of not getting malaria (¬M) is 1. This is different from the .95 probability of the class {H & I}, so it would seem that our explanation (B) is inadequate. However, the law that gets us the probability of ¬M given {H & I & ¬M} is a theorem of probability, so we need not worry. If it weren't for this final unless-clause in (RMS), the only things that would ever count as I-S explanations would be trivially circular deductive arguments, which would not get us too far when it comes to explanation.

The Statistical Relevance Model

Hempel's requirement of maximal specificity leads naturally into our next model, which is Wesley Salmon's Statistical Relevance (S-R) model of scientific explanation. Basically Salmon takes maximal specificity (as given by Hempel) to be *the* important aspect of (probabilistic) scientific explanation. After all, our explanations in this matter are statistical, so if we could find the maximally specific reference class for the explanandum, we would properly understand its probability of occurring. This, says

[9] Or that are believed to be the case

Salmon, is the proper aim of an explanation: to understand the probabilities that the causes involved confer upon their effects. More we cannot ask of any model of explanation.

The notion of 'cause' came up in the last paragraph. Causation, as anyone who has read Hume will know, is by no means a simple or easily understood subject, and many philosophers have been very sceptical about their use in explanation. In fact, Hempel's theories of explanation were designed, in part, to eliminate talk of causes in our scientific discourse. Salmon's model, as we shall see, is an attempt to bring back our common sense notion that causes are important to science and scientific discourse. He writes:

> To untutored common sense, and to many scientists uncorrupted by philosophical training, it is evident that causality plays a central role in scientific explanation. An appropriate answer to an explanation-seeking why-question normally begins with the word "because," and causal involvements of the answer are usually not hard to find[10]

He thinks the best way to discover what the causes of a particular event are is to analyze the reference classes into which that event falls. So an explanatory account answers the question "Why does this x which is a member of A have the property B?" So, for example, "Why did this man with a streptococcus infection recover quickly?" Then the class A must be partitioned into a number of subclasses that are homogenous with respect to B. For a class to be homogenous with regard to property B, all of its members must have exactly the same probability of having property B. So, the class of

---

[10] Wesley Salmon. "Statistical Explanation and Causality." p. 80.

people who have streptococcus infections is not homogenous with respect to quick recovery since it can be subdivided (partitioned) into the class of people who have been treated with penicillin and the class of people who have not, and the probabilities of recovering quickly vary between the members of these groups. But the class of people treated with penicillin is not homogenous either, since it can be partitioned into the class of those who have penicillin-resistant strains of the disease and the class of those who don't, and the probabilities involved are different here. Once all the relevant partitions have been made, we can explain why x had property B by citing the homogenous subclass of A that x is in (x can only be a member of one homogenous subclass – it follows from the definition of a homogenous subclass that they cannot overlap). So, in the case of our man with the streptococcus infection, we might say that he recovered quickly because he was treated with penicillin and had a non-penicillin-resistant strain of the infection. As Salmon puts it, a bit more formally:

An explanation of the fact that x, a member of A, is a member of B would go as follows:

$P(B \mid (A \,\&\, C_1)) = p_1$
$P(B \mid (A \,\&\, C_2)) = p_2$
.
.
.
$P(B \mid (A \,\&\, C_n)) = p_n$

where

$(A \,\&\, C_1)$, $(A \,\&\, C_2,)$, …, $(A \,\&\, C_n)$ is a homogenous partition of A with respect to B,
$p_i = p_j$ only if $i = j$, and
$x \in (A \,\&\, C_k)$[11]

[11] Ibid, 75 (logical notation changed in a meaning-preserving way)

So, to apply this to our streptococcus case:

P(quick recovery | no treatment) = .1
P(quick recovery | penicillin treatment and penicillin resistance) = .2
P(quick recovery | penicillin treatment and no resistance) = .6

These three subclasses represent a homogenous partition of streptococcus infections with respect to quick recovery,
no two classes have the same probability, and
the man in question is a member of the subclass 'penicillin treatment and no resistance'

So here we have a good S-R explanation. Of course, in any real situation the classes and subclasses would be far more complicated, but this gives the basic idea.

What is most interesting about this explanatory account in opposition to the two previous ones is that there is nothing requiring the explanandum to be a member of a subclass with *high* probability of B, or even a subclass that has a high probability of B relative to the others. In other words, if we encounter a second man who has not been treated but has nevertheless recovered quickly, we would simply say that he was a member of the subclass 'no treatment', and apart from that, the explanation would be identical. As long as we know all there is to know about the man's probability of recovering quickly, says Salmon, what would be left unexplained?

When an explanation (as herein explained) has been provided, we know exactly how to regard any A with respect to the property B. We know which ones to bet on, which to bet against, and at what odds. We know precisely what degree of expectation is rational. We know how to face uncertainty about an A's being a B in the most reasonable, practical, and efficient way. We know every factor that is relevant to an A having property B. We know exactly the weight that should have been attached to the prediction that this A will be a B. We know all of the regularities (universal or statistical) that are relevant to our original question. What more could one ask of an explanation?[12]

[12] Ibid. p. 77.

One result of the possibility of explaining low-probability events is that an S-R explanation must not be regarded as an *argument* for its conclusion. Explanation goes through just fine in this case in the absence of inference. The intuitive benefits of this point will be covered more deeply in my chapter "Explanation vs. Inference."

The Deductive Nomological Probabilistic Model

Another influential model of probabilistic explanation is given by Peter Railton in "A Deductive Nomological Model of Probabilistic Explanation."[13] Here he presents, as one might guess, a model of explanation that is very much like Hempel's D-N model, but that allows for probabilistic explanations, hence it is called the D-N-P model.

The basic idea behind Railton's model is that despite our inability to *infer* (deductively or otherwise) unlikely events (such as the decay of a uranium 238 atom in a given day) from the facts about the world, we *are* able to infer the probability they have of occurring. This is a well-known procedure (probabilistic deduction), and is easy enough to perform with only basic logic and basic probability calculus[14]:

1/2 of the marbles in the urn are black
<u>This is a marble drawn from the urn</u>
∴ The probability is 1/2 that this marble is black

or, in symbols…

[13] Peter Railton. "A Deductive Nomological Model of Probabilistic Explanation."
[14] Assuming a single-case propensity view of probability. This is actually a very controversial point that I will not be able to address adequately in this paper. It is not, however, very important to the theory that I endorse what sort of probability one uses, and since I will later give other reasons for rejecting the D-N-P model, I will not challenge Railton on this point.

$\forall x(Ux \supset P(Bx) = .5)$
$\underline{Ua}$
$P(Ba) = .5$

Railton takes basically this form of argument as the backbone of his explanatory account, with a couple of restrictions.

First, like in a D-N explanation, there must be a true law in the premises of the argument that is derived from our scientific theory about the event in question. This law will be much like the laws used in the D-N model, and will be necessary for the derivation of the explanandum. The only difference in the nature of this law and of a D-N law is that it need not be of universal strength. Hence, "All nuclei of U^{238} have probability ($1 - \exp(-\lambda_{238} \% \theta)$) to emit an alpha-particle during any interval of length θ, unless subjected to environmental radiation" is a perfectly acceptable law on the D-N-P account.

Second, there must be, after the conclusion of each D-N-P argument, a parenthetical addendum to the effect that the explanandum event did in fact occur. This might seem like an odd requirement, but remember Railton claims that we cannot give arguments for certain explanantia, we can only give arguments for the probability that they had of occurring. Furthermore, Railton, with Salmon, would agree that once we know "all the relevant factors" that went into an event's occurrence, we have an adequate explanation of that event. All that is left is to say is that the event did in fact occur (probably or improbably), and so Railton adds this as a mere parenthetical aspect of the explanation[15]. So here is the D-N-P model:

$\forall x(Fx \supset P(Gx) = n)$
$\underline{Fa\qquad\qquad\qquad}$
$P(Ga) = n$

$(Ga/\neg Ga)$

and Railton's main example of its application:

(a) All nuclei of U^{238} have probability $(1 - \exp(-\lambda_{238} \% \theta))$ to emit an alpha-particle during any interval of length θ, unless subjected to environmental radiation.
(b) u was a nucleus of U^{238} at time t_0, and was subjected to no environmental radiation before or during the interval $t_0 - (t_0 + \theta)$)
(c) u had probability $(1 - \exp(-\lambda_{238} \% \theta))$ to emit an alpha-particle during the interval $t_0 - (t_0 + \theta)$)

(u did in fact emit an alpha-particle during the interval in question)[16]

One thing that Railton's model has no need for is a requirement of maximal specificity. This is because the law in the premises must meet all the requirements of a law in a D-N explanation. In particular, the property that it predicates in the consequent must be true of *everything* that meets the description in the antecedent. If, therefore, the law in question is not maximally specific (that is, if the reference class of the antecedent can be relevantly partitioned with respect to the consequent), it is simply false.

---

[15] Compare again this view with Salmon's. In the S-R model we set up an elaborate and complicated system of classes and subclasses, but in the end, the explanandum event only gets mentioned with regard to where it falls in these classes. It does not matter what class the event falls into, provided only that it fall into one of them. Similarly, in the D-N-P model, we set up the probabilities involved with x, and then we say whether x occurred or not. The explanation would be a good one whatever happens to be the case with x.
[16] Ibid. 125

This lack of an RMS requirement is a considerable advantage to the view, allowing it to avoid some of the more serious problems that will be raised later with regard to I-S and S-R explanations.

Causal Explanation

Going one step beyond Salmon, David Lewis has developed a theory of explanation that depends *entirely* on the notion of causation. In a paper appropriately titled "Causal Explanation,"[17] Lewis gives a very simple account of this theory:

> Here is my main thesis: to explain an event is to provide some information about its causal history.
>
> In an act of explaining, someone who is in possession of some information about the casual history of some event—explanatory information, I shall call it—tries to convey it to someone else. Normally, to someone who is thought not to possess it already, but there are exceptions: examination answers and the like. Afterward, if the recipient understands and believes what he is told, he too will possess the information. The why-question concerning a particular event is a request for explanatory information, and hence a request that an act of explanation be performed. [18]

This account, unlike any of the other ones mentioned thus far, deals explicitly with causation (remember, even Salmon spoke *explicitly* only of statistical relevance in the model, which he takes to be an indication of causation), which, I take it, is necessary condition of any satisfactory model of explanation (see Salmon's quote on page 13 and my chapter "Explanation vs. Inference").

[17] David Lewis. "Causal Explanation,"
[18] Ibid. 217-18

Lewis's theory is itself based on Sylvain Bromberger's theory of explanations as events that can be described by "A explains X to B" (where A and B are both people and X is the explanandum-event)[19], and like Bromberger's it is far more general than it might at first seem. For instance, an explanation can exist independently of anybody's knowledge of it, as the "causal information" that one would provide if one were to explain a particular event. This is how a scientist can meaningfully say "there is likely a theory out there that explains this phenomenon, but nobody knows it yet." In fact, the very act of explaining can be performed not just by a person, but also by a hypothesis, by certain premises, and by evidence. On most accounts of causation (on Lewis's, at least), this theory would imply that there are a great number (perhaps infinitely many) of correct explanations of any particular event. What determines whether a given explanation is a good or useful one is just a matter of the speaker's context. For example, say a baseball strikes and breaks a window. In this case a scientist studying the nature of the baseball's flight could truly say that its speed and weight explain the window's breaking, while a scientist studying the strength of the window could truly say that its thinness and fragility explain the same event. In fact the same person might, in different contexts, regard different causes as important. The doctor whose patient develops a rash after his penicillin injection might tell the patient that "it happened because of the injection," and tell his allergist acquaintance that "it happened because of the patient's penicillin allergy." In neither of these cases, however, would the

[19] Bromberger, Sylvain. "An approach to Explanation."

explanation "because the big bang occurred" be a *good* one, despite the fact that it provides causal information and is therefore and explanation. Each of the contexts forbids mention of such general and remote causes.

Lewis's theory is a very loose and general account indeed. The advantages of the looseness of this theory are obvious. Lewis is easily able to account for explanations across a broad range of scientific disciplines, and across a broad range of theories within each discipline. Furthermore, as I mentioned briefly above, the causal nature of this account gives it an intuitive plausibility. If we can cite the causal process which brought about a particular event, what more could possibly be asked?

CHAPTER III

SOME PROBLEMS WITH THE STANDARD MODELS

Each of the above models has its distinct advantages, which I have tried to highlight, but each has considerable drawbacks as well. I will show some of these drawbacks in the hope that through understanding them we can gain a clear idea of what exactly is required of an adequate account of scientific explanation. That is, any satisfactory account of scientific explanation must avoid all the pitfalls that these models encounter, without having any serious problems of its own.

The Problems

Problems for Deductive-Nomological Explanations

One deficiency of the Deductive Nomological model is that it fundamentally involves laws, and laws are very tricky things to understand. They are generally taken to be logically general empirical statements that contain no singular terms, and that can (in principle) be subjected to unlimited tests. This would rule out any statements that refer to particular people, places or things, as well as any statements that referred to particular times. It would rule out such *logical* truths as the *law* of non-contradiction and the *law* of the excluded middle term[20]. What it doesn't do is differentiate between these two statements:

[20] Which is, of course, a good thing. Unless we have a Fregean conception of logic as a *scientific* inquiry, we would not want such statements to count as *scientific* laws.

A: "No solid sphere of pure uranium weighs more than 100,000 pounds."
B: "No solid sphere of pure gold weighs more than 100,000 pounds."

A and B both fit the definition given above of lawlike sentences. Both are presumably true, but A represents a *law* while B does not. Realistically, we can understand why A is lawlike and B is not. It is because any mass of uranium would reach critical mass and explode long before it got to 100,000 pounds, while gold has no such instability. But to give a strict logical formulation of why the first is a law and the second is not is by no means an easy task[21], and we will need such a formulation before we can confidently apply the D-N model.

Another problem with the D-N model is simply that it is deductive. Many philosophers (John Stewart Mill being one of the most vocal) hold that deductive inferences are epistemically useless since there is no information imparted by the conclusion that was not already known in the premises. If I say to someone '$(P \supset Q) \supset P$', on this view, then that person has also been told that 'P' (which follows logically), whether or not they realize it. If this view of deduction were true, scientific explanations of the D-N sort would indeed be questionable. What sort of explanation *in principle imparts* no new knowledge on the hearer? In fact, I think that this view of deduction is mistaken. Would we say that we have no more information regarding the truth of Fermat's last theorem now than we did twenty years ago? No, and yet the proof that gets us there[22] is entirely deductive. To take my earlier example, when I first prove 'P' from

[21] Such formulations usually involve counterfactuals and other sorts of modal discourse. But with such logical apparatus it can be very difficult to provide consistent, non-circular reasons why A is a law and B is not.
[22] Which I will not be including in this paper for a multitude of reasons that will, I hope, be obvious.

'$(P \supset Q) \supset P$', I have added something new to my body of knowledge, and traversed an "epistemic gap."

Despite the usefulness of deduction however, it is somewhat odd to imagine that science, in explaining the world, is just doing logic or mathematics. While it is certainly appropriate that scientific explanations *involve* mathematics and logic, it is a more dubious claim to say that they just *are* mathematics and logic. I am sceptical of any sort of explanation that does not essentially involve reference to the process through which the explanandum took place, or that merely explains by pointing out that "things are always (or usually) this way." I do not regard the following inference as explanatorily helpful:

All iron rods conduct electricity
x is an iron rod
∴ x conducts electricity

It leaves me completely unsatisfied as to *why* my iron rod conducts electricity. This point will be addressed in detail in my chapter "Explanation vs. Inference," so I will not discuss it further here. Suffice it to say, I consider it a fatal objection to the D-N model, and to a number of the models that follow.

Problems for Inductive-Statistical Explanations

The I-S model of explanation has many of the same problems as its deductive counterpart. I-S explanations, like their D-N counterparts do not seem to add to my *understanding* of the event in question. Take the previous example of an iron rod conducting electricity and imagine it as an I-S explanation:

> Every iron rod has a probability r of conducting electricity.
> x is an iron rod.
> ==[r]
> x conducts electricity.
>
> Where r is close to 1 and both premises are contained in K, and the class of iron rods is maximally specific.

It just does not seem that this I-S explanation adds anything to my understanding of *why* my iron rod conducts electricity. Just knowing that most rods have the property in question leaves my “why?’ question unanswered. Construing explanations as inferences, deductive or inductive, is problematic, as my chapter “Explanation vs. Inference” will make clear. The more specific problems involved with I-S explanation, however, involve its epistemic relativization and the requirement of maximal specificity.

The severity of the problems posed by the epistemic relativization of I-S explanations under RMS is argued for by J.A. Coffa in "Hempel's Ambiguity."[23] The basic argument begins with the fact that RMS implies that I-S explanations must be relativized to a particular person's epistemic situation, and that there are therefore no independently "true" I-S argument (which is a characterization that Hempel would agree with). When the fundamental subject matter at hand is a person's epistemic situation, the only sorts of useful results we can get are *confirmations* of certain beliefs (for that

[23] Alberto Coffa. "Hempel's Ambiguity,"

person) relative to his or her other beliefs. This notion of confirmation is used all the time in both scientific and non-scientific discourse. For instance, in a D-N model, we can speak of the explanation being *confirmed* by a person's evidence for it, which means that they will come to believe it (the explanation) to be true. And a detective might say that certain evidence confirms her in her belief that a particular person committed a particular murder, which means she believes (more or less strongly) that it is true that the person committed the murder. In both cases we think of confirmation as strengthening a person's belief that a particular claim is true. But I-S explanations are importantly disanalogous. They are (by definition) not the sorts of things that can be true, so nothing could adequately confirm us in the belief that they are true. What then *is* confirmed in an I-S explanation? It cannot be the I-S explanation, as we have said. It cannot be the explanandum, since we knew that to be true at the outset. In cannot be the explanans, since the explanans is (by definition) what does the explaining, and for a statement to explain itself is viciously circular. What then does the I-S argument confirm? In the end, Coffa leaves this question open, and until it gets an answer (and I don't have one) it does not seem very reasonable to accept the I-S model of explanation.

The second problem posed by the RMS is one that will affect not only the I-S model, but also Salmon's S-R model (which I shall get to in a moment). If an explanation, to be acceptable, must refer only to the most specific reference class of the cause in question relative to the effect in question, our reference classes seem to get very small very quickly. After all, with regard to John's not getting malaria, we cannot put him in the class {H & I & ¬M} since the resulting probability of ¬M would be a theorem of probability. We can however, put him into the class {H & I & J}, where J contains information that pertains to John in particular. The result would be a new reference class in which the probability of ¬M is 1 (since John didn't get Malaria), not .95, and so this would have to be the reference class that we use in our I-S explanation[24]. It might be objected that a reference class ought to omit any particular or singular terms, but we could get the same results if we just describe *enough* of John's more general features, such as being 6 feet tall, being 21 years of age, having black hair, having type O- blood, and so on. Eventually we would get a description that *only* John fits that contains no singular terms. It might be further objected that the RMS requirement is relativized to a knowledge-situation, but this does not help, since all of these descriptions of John might well be a part of that situation (if one were his doctor, say), and we would therefore be *required* to consider the extent to which they affect the probability of getting malaria, which they do. So if we were always required to make the reference class as specific as possible, we would usually wind up with "statistical" arguments that confer the probability of 1 upon their conclusion, which is a result that is to be avoided, if at all possible.

Problems for Statistical-Relevance Explanations

One of the main problems with Salmon's S-R model involves causation. Salmon feels that causation is fundamentally important (as quoted on page 13), but there is no

---

[24] It is important to note that it is *not* a theorem of probability that the probability of ¬Mx given that

mention of causes in the formal explanatory account. If Salmon were indeed silent on the matter of how causation affects his model, this would be a rather serious inconsistency, but it is quite clear that, far from ignoring causation, he merely treats it as something that can be adequately understood in terms of statistical relevance. In the quote on page 15, Salmon mentions that on his account, we know "every factor that is relevant to an A having property B," and wonders "what more could one ask of an explanation?" Well, since the rhetorical answer to this rhetorical question is "nothing," and since Salmon himself explicitly requires explanations to cite causes, we can infer that he takes statistical relevance relations to be, in some sense, causal. Clearly, Salmon believes that anything that is really causal with regard to B will show up as part of a reference class description. So in the above explanation I think it is intuitively reasonable to regard being infected, being treated with penicillin, and having a non-resistant strain as causes of a quick recovery and, as it happens, each of these plays a part in the description of at least one homogenous subclass. Presumably for any A to cause any B, the A must have an effect on the probability of the B. If this is true, then A *must* have a place in the partitions within our model. So, problem solved—causation implies statistical relevance, so causation is involved with the model.

This solution, however, leaves open an important question, "does statistical relevance imply causation? I do not think that it does. Having a penicillin-resistant strain is relevant to the probability of a quick recovery, but we would never say that it *caused* a quick recovery. On the contrary, we would say that a quick recovery occurred

---

$x \in \{I \& H \& J\}$ is 1, so long as we keep $\neg Mx$ out of the description in Jx.

despite the resistance of the infection. But if statistical relevance does not imply causation, how can we tell, using the model, which the causal factors are, or indeed *whether or not there are any causal factors.*

The only answer I can imagine is that we might regard those factors that raise the probability of B over the probability in the next-class-up could be causes (of B), and those that lower the probability of B over the probability in the next-class-up are non-causal (or are counteracting causes) with respect to B. This answer is good as far as it goes, but it runs into a problem when we try to determine whether "having the streptococcus infection" is a cause of "quick recovery." In all but the most general classes of events, the class A (the most general class in our explanation) would itself be a subdivision of some broader class C. If we allow for an analysis of A's causal role in terms of a class C of which it is a subdivision (say, the class of all humans), then we will know whether or not A causes B, but we will be left with the further question of whether C causes B, and the analysis would go on indefinitely[25]. Such an indefinite procedure is a problem for Salmon because he requires all the causally relevant factors to be included if the explanation is to be a good one. Another option is to say that the analysis of causes *must stop at* A, but this is too arbitrary a restriction on a causal analysis to be taken seriously. If our explanatory model forbids us from asking whether or not being infected with streptococcus can be a cause of a quick recovery, then so much the worse for our explanatory model. Furthermore, such a restriction would allow explanations

[25] Unless of course we have in mind a definite stopping point, like "the class of all objects, properties and events of any sort," but such a stopping point is rarely, if ever, to be reached in our statistical subdivisions.

that cite no causes whatsoever. For instance, a man who has a streptococcus infection and is not treated, but who recovers quickly. If we cannot inquire about the causal significance of "having a streptococcus infection," then we can have an explanation of the quick recovery that cites no causes whatsoever (or only counteracting causes), and this is problematic, given Salmon's own statements regarding the importance of causation to explanation.

Causation aside, the statistical relevance model is basically a non-epistemic version of the requirement of maximal specificity expanded and made into its own model of explanation. Fortunately, since the S-R model is not relativized to any particular knowledge-situation, Salmon can avoid the objections that Coffa raises to Hempel's RMS. Unfortunately, it is still a version of RMS, and this means that it will have the problem of determining when certain description are maximally specific and the problem of the overly-specific reference class mentioned in connection with John's malaria. It is incredibly difficult to think of even a simple real-life example where all the relevant statistical information is understood (the streptococcus example is, of course, grossly oversimplified)[26], and even if one could, we would still have the question of why the most specific reference class is not the one that describes just the object or situation in question, thus rendering all the probabilities involved either 0 or 1.

---

[26] This objection (that all the relevant statistical information could never be known), which is problematic but not fatal to Salmon's model, is one that Hempel is able to avoid by relativizing the explanation to an epistemic situation. It is not necessary that *all* the statistically relevant information be known for Hempel, merely that we apply what we do know properly. Of course, this opens him up to Coffa's objection.

Problems for Deductive-Nomological-Probabilistic Explanations

Railton's D-N-P model is, as the name suggests, basically a version of Hempel's D-N explanation. This means that it is open to many of the same objections.

First, Railton, like Hempel, makes important use of laws, and therefore takes on all the difficulties that they present. The fact that Railton's laws can be probabilistic does not save him from counterexamples like the solid sphere of gold law. After all, universal laws can be given a probabilistic form as follows:

A': The probability that any solid sphere of pure uranium weighs more than 100,000 pounds is zero.
B': The probability that any solid sphere of pure gold weighs more than 100,000 pounds is zero.

Or, in symbols:

A' $\forall x(Ux \supset P(Wx) = 0)$
B': $\forall x(Gx \supset P(Wx) = 0)$

It does indeed seem that A' is a lawlike statement and that B' is not, but it is by no means clear (just as it was not clear on the D-N account) how we are to show that difference within the formal apparatus with which we are provided.

The other problem that the D-N and D-N-P models share is that they are both argument-based (inferential) theories of explanation, and therefore, I believe, leave our "why?' question unanswered. Admittedly, the D-N-P model differs somewhat from the D-N model in that it does not deduce the *fact* that the explanandum occurred, but merely the *probability* that it had of occurring. But being told that the probability of x was n does not seem to help us with the question "*why* did x occur?" We are left wanting more of an explanation of what went on to bring about x, what process led to x, what *caused* x.

Problems for Causal Explanation

The primary (and only truly significant) disadvantage to Lewis' theory is its looseness. I mentioned above that this was a particular *advantage* of his account, but it turns out that its greatest strength is also its greatest weakness. For one thing, such a broad analysis of explanation as this one hardly seems like an analysis at all. Lewis explicitly states that he feels it is unnecessary (and perhaps impossible) to offer necessary and sufficient conditions on explanations, since our intuitions change frequently depending on both the changing contexts of the explanandum and depending on our own changing contexts. But necessary and sufficient conditions *just are* how we analyze terms, and so Lewis must be giving us something a little less exact than an analysis, like "this is how we ordinarily use the word 'explanation'."

I think that Lewis would agree, saying that an analysis is indeed impossible, but that it is also unnecessary. But might it not be perfectly appropriate to give restrictions on what may count as a *scientific* explanation, even if our ordinary discourse is too complicated for such restrictions? Wouldn't it be proper to formalize and regulate this particular procedure of science so that we have a clear direction to our inquiry, and a good understanding of when success might have been attained? In the end, Lewis has offered us a useful description of our ordinary explanatory discourse, but has not given us a model precise enough to apply to the physical sciences.

Conclusion

So where does all this lead us in terms of developing a more correct theory of scientific explanation? Well, we will want a model that has all or most of the advantages of the models already given and few or none of the disadvantages.

So we will want a reasonably rigid formal model that essentially involves causation and is able to handle probabilistic explanations. It ought not involve laws, nor should it have any requirement of maximal specificity. It should make our scientific explanations objectively true (that is, not relativized to any epistemic situation) and empirically verifiable. Finally, and perhaps most importantly, our explanations should not be inferential, but rather should really tell us *why* the event happens (not just that it was to be expected).

I contend that Paul Humphreys has presented just such a model, entitled "Aleatory Explanation." In the next chapter I will outline his view and show its many advantages to those given here. In the end, I believe that only a few minor revisions must be made in order to render Humphreys' account *the* correct model of scientific explanation.

CHAPETR IV

ALEATORY EXPLANATIONS

In the last chapter, I outlined several of the major models of scientific explanation in order to both gain an understanding of the subject and to identify some of the major difficulties with these models and where they lead us. In this chapter, I will be discussing Paul Humphreys' view of "aleatory explanation," which I believe avoids the pitfalls of these other views. As a result, it is a much more useful and realistic account of scientific explanation than anything else in the literature.

Humphreys' General Account of Explanation

Paul Humphreys, in "Aleatory Explanations,"[27] constructs an explanatory model that is based entirely on statements of causality. It is very much like Lewis's theory of causal explanation in that it takes an explanation to be something that provides information about a particular event's causal history. But he differs from Lewis in three important ways. First, Humphreys is far more restrictive in his model, and not just any statement of causes will do. Second, context is not relevant to determining the quality of an explanation. Third, his explanatory account makes a distinction between counteracting causes and contributing causes. It seems to me that Humphrey's model is the most intuitive and least problematic of all those mentioned here, and my main project

[27] Paul Humphreys. "Aleatory Explanations."

throughout the rest of the thesis will be to argue in support of it and examine solutions to the few problems that it does have.

Preliminaries out of the way, the basic (canonical) form of an aleatory explanation is:

A because Φ, despite Ψ.

Where Φ is a non-empty set of contributing causes, Ψ is a set, possibly empty, of contributing causes, and A is a sentence describing what is to be explained.[28]

The notion of what exactly counts as a cause is by no means an easy issue, as Humphreys will agree, and it will of course be necessary to shed at least some light on it to make sense of this model. First, he intends his model to be effective for both deterministic and probabilistic causation. This goes back to a point I made in the introduction regarding the adequacy of any explanatory model; it should not, in itself, determine the validity or invalidity of any particular scientific claim. So since the debate between determinism and indeterminism (and the related debate between deterministic and indeterministic causation) is still a very live and relevant one, it would be inappropriate for our model of explanation to assume either side of the issue. But though he remains neutral on the determinism/indeterminism issue, Humphreys does give us a general outline of what counts as a cause according to his model. This is necessary because, 1) he needs to defend the idea that there can be multiple causes of an event, 2) he must also defend the idea that there can be both contributing and

[28] Ibid. 227

counteracting causes, and 3) he needs to show how his model can be applied (or is already applied) to explanations within the scientific community.

Causation

The first thing that must be defined within Humphreys' bare-bones account of causation is an 'event'.

> An event is the possession of, or change in, a property of a system on a given occasion (trial). Events are thus taken as concrete, specific entities, actual instantiations of or changes in worldly properties of a system, these properties being possessed by specific structures, themselves a part of the world, with these structures persisting through the change in properties which constitute an event.[29]

Humphreys will only be concerned with causation involving events as defined above. In other words, both the effect and the cause in question must fit the above definition in order for the explanation to be acceptable. While this might seem restrictive, I think the definition of an event is broad enough to allow for any sort of causation that would be appropriate within scientific discourse.

One thing Humphreys' account of explanation (which allows for a multiplicity of causes) must argue against is the notion that the only real causes of an event are *singular* causes. Singular causes are causes that, regardless of other circumstances, always produce the same effects. So if X is the cause of Y, then Y will occur whenever X does, no matter what other circumstances obtain. This idea has a certain intuitive appeal, but

[29] Paul Humphreys. "Scientific Explanation: The Causes, Some of the Causes, and Nothing but the Causes." p. 289.

unfortunately, the world is just far too chaotic, and few, if any, such regularities exist. Humphreys considers the objection that when experimental controls are place on nature (as in a laboratory), certain regularities do become apparent, thus supporting the idea that there could be singular causes, but dismisses it as too limiting to the project of science. As the goal of science is to explain the world as a whole, not just the world of the laboratory, it is unrealistic to confine one's notion of causation in such a way. Hence, when we discuss causes, we must have a looser definition in mind than *singular* causation, otherwise we would always (or almost always) be incorrect.

Another apparent difficulty for Humphreys' view of causation is the notion of a *counteracting* cause. A counteracting cause is, simply, a cause that acts *against* the effect. The reason this seems like a problem is that we earlier allowed that the causes in question be either deterministic or indeterministic. The indeterministic case seems unproblematic, with counteracting causes merely lowering the probability of the event, but not ruling it out, but in the deterministic case, it would seem odd to think of anything working *against* the effect without precluding it altogether. To understand how deterministic explanations can have counteracting causes, it is necessary to draw a distinction between purely qualitative explanations and quantitative explanations. Pretty much any event can be understood (and explained) in terms of both its pure qualities and in terms of the quantitative values of those properties. Humphreys gives the case of a room that is both heated and air conditioned. Over the course of one hour, the heater pumps 10,000 Btu into the room and the air conditioner pumps 5000 Btu *out* of the room, and the net rise in the temperature of the room is 5˚C. When we ask, "why did the room get 5˚C hotter?" we could be asking why the temperature increased, or we could be asking why it increased *by* 5˚C. Both would be legitimate explanation-seeking "why?" questions, and both are proper within scientific discourse, but the answer would be different in each case. In the quantitative case, our previous intuition would be correct, and there would be no counteracting causes. We would just say "The room got 5˚C hotter because of the combined actions of the heater and the air conditioner." This is because, had either of them been absent from the situation, the temperature would not have risen *by 5˚C,* but by some other amount (or not at all). In the purely qualitative case, however, we would say the temperature rose "Because of the action of the heater, despite the action of the air conditioner." In this case, the effect (the temperature rising) would have been accomplished without the air conditioner, and in fact, it worked *against* the heating of the room, so it is reasonable to speak of it as a counteracting cause. What this example demonstrates is that while it is often appropriate to speak of counteracting causes in qualitative deterministic explanations, it is rarely, if ever, appropriate to speak of counteracting causes in a quantitative deterministic explanation. This is, however, unproblematic for Humphreys' view, since the set Ψ of counteracting causes can be the empty set.

With regard to how his model can be applied to actual scientific explanations involving causation, the preceding example is a very good case. Whatever scientists might have in mind when they speak of causation, it would seem that the causes mentioned in the heater/air-conditioner case must be included. Another example Humphreys gives is that of a car skidding off the road. We might say that it skidded off the road because it was traveling excessively fast, the road had a steep slope, and there was sand at the intersection, despite the fact that the brakes were good and the tires were properly inflated. Here we have a number of causes (counteracting and contributing) put into the canonical form of an aleatory explanation. Each in definable in purely physical terms (velocity, gravity, inertia, friction, and such), and each of them either adds to or subtracts from the probability that the car will skid off the road. Surely there were other factors involved (weather conditions, driver alertness, etc.), but despite that, it would appear that we have at least a partial explanation of the event in question in terms of quantifiable and scientifically understandable causes. This is, of course, just an example of how Humphreys' account of causal explanation matches up to our physical and scientific discourse regarding causation.

One might worry that there is no stipulation that the causes involved in the explanation be explicitly of this physically quantifiable sort. Hempel, for instance, would probably worry that this opens up the door to all sorts of pseudoscientific explanations of events (astrology, crystals, and the like) [30]. But these fears are misguided. What *is* stipulated is that any cause of a particular event must either raise or lower the probability of that event (for contributing and counteracting causes, respectively). If the position of the stars and the use of crystals have no impact upon a particular event, then it would be inappropriate to include such information in our explanatory account. If it turns out that stars and crystals *do* impact the physical events in question (such as cars skidding off the road or rooms becoming hotter), then we were

[30] This is one of the motivations behind the symmetry thesis. See my chapter on contrastive explanations.

wrong to think of such things as pseudoscience in the first place, and we *should* include such information in our explanations when relevant. The point is that it is not up to our theory of explanation alone to determine the nature of causation or the proper place for astrologers. Presumably the scientists can be trusted to put the pseudoscientists in their place through rigorous experimentation and theorizing.

However, one might point out that since the stars exert gravitational force upon everything in the universe (presumably), they must have *some* effect upon everything that happens. One might further argue that given the conditions on what counts as a cause, this gravitational impact should be considered seriously, and the stars therefore *do* explain events on earth, according to Humphreys. As the stars clearly do *not* explain such events, we seem to have a reductio on Humphreys' view. And more generally, since most events impact most other events, it would seem that our explanations would become hopelessly long and complicated, and would involve many apparently irrelevant statements of very weak causal connections, which would, of course be problematic. In order to respond to this objection, I will have to discuss the *incompleteness* of scientific explanations.

Humphreys takes incompleteness to be a normal feature of scientific explanations. Since our knowledge of the physical world is never quite complete, neither are our explanations. But even if our explanations are incomplete, they can still be correct. Humphrey's gives the following case:

> Consider how we would respond to a request for an explanation of the increase in angular momentum of the earth over the conserved value. The explanation would be "because of the precession of the moon's orbital plane, the nonuniformity of the sun's

> gravitational field, and the action of thermodynamical tides, despite the slowing effect of tidal friction."…There is no pretense that all causal factors affecting the angular momentum of the earth have been cited. The omissions are not due to the scientist selecting those factors which interest his listener, or to being constrained by the form of the request for an explanation. It is because there are many influences on the earth's rotation, most of which are as yet unknown. The geophysicist knows that there exist these unknown causal factors, yet the factors cited do provide an explanation, however incomplete, of the explanandum. Nevertheless, the explanation given is true.[31]

So in the case of our astrological example, it would not be necessary, on Humphrey's model, to provide any information whatsoever about the gravitational impact that the stars have on earth-bound objects. If such information (properly sorted into contributing and counteracting causes) were included, it would not hurt the explanation (in fact, it would help a tiny, tiny bit), but it is not necessary in order for the explanation to be correct. Further, in response to the more general objection that aleatory explanations would become needlessly long and complicated, there is no reason to think this is the case. Their length and complexity would (or should) merely reflect the explainer's level of knowledge and expertise, which seems appropriate.

Quality

Another important feature of aleatory explanations is that they have varying degrees of quality. That is, there can be better and worse explanations of a particular event, all of which fit the canonical form, and all of which are true. This is a feature that many of the other models I discussed do not have. The D-N, I-S, S-R, and D-N-P models all have one and only one true explanation of a given event, and therefore there

[31] Ibid. 288

can be no diversity in the quality of explanations. Any explanation is either true or false, with no middle ground[32]. Lewis' theory of causal explanations does allow for better and worse explanations, but only relative to a given speaker/listener context. This, as I have pointed out, does not seem quite objective enough for the project of scientific inquiry. Humphreys' conditions on the quality of an explanation are as follows. For any two true explanations of a particular event Y (keeping in mind the canonical form of 'Y because Φ, despite Ψ'), "If Φ _ Φ' and Ψ = Ψ', then the explanation of Y by Φ' is superior to that given by Φ. If Φ = Φ' and Ψ _ Ψ' then again Φ' gives a superior explanation, in the sense that it is more complete."[33]

So, simply put[34], explanation A is better than explanation B if either A cites more contributing causes and the same counteracting causes as B, or A cites more counteracting causes and the same contributing causes as B. I imagine (though Humphreys himself never explicitly points this out) that Humphreys would also agree that when A gives both more counteracting causes *and* more contributing causes than B, A is the better explanation. Of course, even with this addition, these are merely sufficient conditions on what constitutes better and worse explanations. Humphreys does not give us any way of telling, for *any* two explanations, whether one is better than

---

[32] Though as I mentioned in the previous chapter, Hempel denies that I-S explanations should ever be seen as true, or even "true within a certain epistemic situation." However, given Coffa's objection, Hempel will need to allow I-S explanations to be true or "good" in some sense, or else the view falls into meaninglessness.

[33] Ibid. 296

[34] And hopefully not too misleadingly put. It should be kept in mind that when I say "A cites more causes than B" I mean "A cites all the causes that B does, plus some that B does not." The 'proper subset' notation that Humphreys uses is really the best way of expressing this, but is not easy to put into simple English.

another. In the case of two explanations that cite *different* causes, with little or no overlap, we are left to our own devices to determine which one is better. This is entirely appropriate. It should not be up to our model of explanation alone to determine whether the car's skid is better explained by "because of its speed, despite its working breaks," or "because of the angle of the turn and the sand on the road, despite its properly inflated tires." It is the physicists job to properly sort out the causes and determine which ones are most important. This sort of inquiry is the proper work of science, and there is no reason why our theory alone should be expected to replace it.

It should also be noted that these conditions on the quality of explanations cannot be used to determine which explanations are "good" or "bad." They can only determine the quality of an explanation relative to another explanation of the same event. This is not to say that there are no good or bad explanations, merely that such a feature of an explanation cannot be determined by analysis of the explanation's form alone.

Though it may be obvious, I must point out that an explanation which incorrectly identifies a contributing cause as a counteracting cause, or a counteracting cause as a contributing cause cannot be considered a "better" or "worse" explanation at all. It is simply *not an explanation*, or it is a false explanation. This just follows from what it is to be an aleatory explanation. An explanation is true iff it meets the canonical form of an explanation, as given above. Since an explanation that includes counteracting causes in the set Φ or contributing causes in the set Ψ is does not fit the canonical form, it is not a true explanation at all, no matter how convincing or clear it might be.

Probability Values

It should be noted that unlike the other models I have covered that allow for indeterminism, Humphreys' does not include any statement of the probability values involved. Hempel's I-S model and Railton's D-N-P model each conferred a probability value on their explanandum, and Salmon's S-R model consisted of very little but a series of probability statements, but here Humphreys has given us a probabilistic model that does not make mention of probability values at any point. Why is this? There are, in fact, two reason for this. One is simple and relates to some points I have already covered. The other is more complicated and stems from Humphreys views of the nature of probabilities and propensities. I will explain each briefly.

The simple problem with including probability values in an explanation is that they will almost invariably make that explanation false. This relates to a problem that Salmon's S-R model had. Basically, the problem is that there is always (or almost always) some information that we do not have, or that we are unaware of the implications of, which will throw off our probability calculations, thus rendering the values in our explanation incorrect. So, to use our previous example, John's chances of contracting malaria given that he has been infected with malaria plasmodium and that he is a heterozygote with regard to the sickle hemoglobin gene are .05. Now, say we put together an explanation of his *not* contracting malaria that cites this value: "John did not contract malaria because the probability of a heterozygote with regards to the sickle hemoglobin gene who has been infected with malaria plasmodium contracting the disease is .05.[35]" Now, imagine that we did not know that John had been taking quinine faithfully for months prior to the infection. This extra true fact about the world lowers the probability that John will contract malaria from .05 to, say, .01. Now, simply because we were missing a piece of information about the world, our explanans is false. As such, it cannot possible hope to explain John's not contracting malaria[36]. Since, as we discussed earlier, completeness is not to be expected in science, it would seem that there will almost always be some little unknown fact that alters the probability value in any explanation somewhat. To take a more extreme case, imagine that, unbeknownst to us, John had drunk one cup more than the average amount of water in the two weeks prior to his infection, and that this lowered his chances of getting malaria from .05 to, say, .0499999. Even such a trivial change in the probability values would make an I-S, S-R, or D-N-P explanation false. As a result, explanations that involve probability values are not practical in science, and if we happened upon a true one, it would almost certainly be by luck.

But consider the first of the above cases from the standpoint of an aleatory explanation. We would say (assuming that we still don't know about the quinine) that John did not contract malaria because he is a heterozygote with regard to the sickle hemoglobin gene, despite the fact that he was infected with malaria plasmodium. This

[35] I realize, of course, that the other models of explanation would not use this linguistic form (which is a variation on the canonical form, except that it involves probability values), but the criticisms should work no matter how you set it up.

[36] Hempel might disagree. His I-S model is relativized to a particular epistemic situation. This means that the counterexample does not work against I-S explanations, since the facts about quinine were not in our body of knowledge K. But as we have seen the difficulties Coffa raised with regard to epistemic relativization, I hardly think that Hempel's view is to be preferred here.

incomplete aleatory explanation is true even though we are missing some information. It would certainly be a *better* explanation if it included the relevant facts about the quinine, but that is not necessary in order to have a partial explanation. Most importantly, because we did not cite any probability values, nothing in the statement is false. As long as we have correctly identified some contributing and counteracting causes, we have given an explanation of the explanandum. So the main reason not to include probability values in aleatory explanations is simply that we want at least some of the explaining done by scientists to be *correct*.

It might be objected that even Humphreys' account can fall victim to the general problem of missing information. For instance, imagine our case with the car skidding off the road. We might say "it skidded because it was traveling too fast and there was sand on the road, despite the fact that the brakes were good." Now imagine that, unbeknownst to us, there was ice on the road. Well, now the presence of sand turns out not to be a contributing cause of the skid, but rather a counteracting cause of it (increasing the sliding friction, rather than decreasing it). So our apparently good aleatory explanation turns out to be false because of some information we didn't have. While it is true that this danger exists with aleatory explanations just as it does with the others, and while this does mean that apparently good aleatory explanations are often incorrect, at least with aleatory explanations, there is some chance of being correct reasonably often. The fact is, anything that would make an aleatory explanation false would make any of the other accounts (I-S, S-R, D-N-P) false as well, but most of the things that make those accounts false would not make an aleatory explanation false. It seems sensible to stick with the view that at least gives us a chance to produce correct explanations, even if it is still difficult and requires careful study and experimentation.

The second problem that Humphreys has with including probability values as a part of explanations stems from his views about exactly how probabilities relate to propensities. A propensity is loosely defined as an indeterministic disposition that a natural system has. The most obvious example of something that is widely believed to involve such a propensity is the decay of a radioactive atom. Our best science at the moment tells us that any given atom of Uranium 238 *might* decay over a particular time, but that it is in principle impossible to tell for sure. This irreducible *chance* is often called the *propensity* of the atom to decay. It seems like there are (or at least might be) such propensities in the natural world, and it seems like our theory of probability ought to be able to handle them. After all, we seem to be able to make sense of expressions like "That atom has a probability of .01 of decaying this year," and it just seems like probability, if it corresponds to anything in the world, must correspond to chances. But unfortunately, the apparatus of the probability calculus has incredible difficulty with this matter. It is usually seen as more fitting mathematically to regard probabilities as long-term frequencies within a system. So in the case of our atom, the frequency interpretation would say "In the long run, .01 of all atoms of this type will decay within any given year." But this is very different from saying *this* atom has a particular chance of decaying. To transfer the frequency interpretation onto propensities in this manner commits the fallacy of division, assuming that the parts are like the whole. So there is a rather serious problem here that has still yet to be resolved. I will not pretend to solve it, but I will give Humphreys' view of the matter, which should shed some light on why he doesn't want probability values in his aleatory explanations.

In "Why Propensities Cannot be Probabilities,"[37] Humphreys puts forth a rather detailed and technical argument that can fortunately be summarized more simply as follows. Probability calculation involves a certain symmetry between the two parts of a conditional probability statement. So if you know the probability that a given draw will produce a black marble given that it came from urn A, and you know the probability that the draw will be from urn A and the probability that it will produce a black marble, then you can deduce the probability that it was drawn from urn A given that it is black. So from the probability value of "B, given A" we can (provided sufficient additional data) deduce that value of "A, given B" in every case of a conditional probability. But consider now a propensity translated into the form of a conditional probability. A simple case might be "This piece of metal has a .5 propensity to emit an electron given that it is exposed to a particular amount of light." But even if we knew the probability that it would be exposed to the light and the probability that it would emit an electron, it would still seem odd to say that we know the propensity it has to be exposed to a particular amount of light, given that it emits an electron. Since the propensity is in some sense causal, it carries with it all the asymmetry of causality. We can say that a window has probability 1 of breaking, given that it is hit by a rock, but it seems almost nonsensical to ask what its probability of having been hit by a rock is, given that it breaks. So Humphreys develops this rather informal argument in more formal terms, and winds up

[37] Paul Humphreys. "Why Propensities Cannot be Probabilities."

with the conclusion that probabilities cannot be propensities. But rather than take the standard approach and try to make our scientific models fit the frequency approach that is so much more appropriate for probability, Humphreys just says, in effect, "so much the worse for the probability calculus if it can't handle propensities." So Humphreys endorses the idea that there are chance phenomena in the world, but denies that the probability calculus is the proper way of dealing with them.

We can see then what his other motivation behind not including probability values in his explanatory account is. He simply doesn't think that probability values (properly understood) have any place in our discussion of the propensities embodied by the physical world. Together with the more obvious fact that requiring probability values in our explanations will virtually always make them false, we can see that there are good reasons to omit the precise probability statements endorsed by other models in favor of Humphreys' account.

The Superiority of the Aleatory Model

Now let me review the conditions of the adequacy of a model of scientific explanations that I developed in the last chapter. If you recall, these conditions were developed by examining some of the shortcomings of the other prominent theories of explanation, and so any explanatory model that satisfies them would be at least somewhat more acceptable than these other models (provided, of course, that the model in question is internally coherent, accords with our intuitions, and accurately describes the procedures of science). I will go through each of these conditions one at a time and determine whether or not Humphreys' model satisfies them.

First, the model ought to be rigid enough to clearly distinguish true from false explanations, and scientific from non-scientific explanations. It is clear that aleatory explanations, unlike Lewis' theory of causal explanations, gives us necessary and sufficient conditions on true explanations, provides us with a somewhat rigid linguistic format, and even gives us some notion of what counts as a cause within his theory. To the extent that his theory might be seen as containing any looseness, it is a looseness implied by the diversity of, and difficulties with actual scientific explanations. For instance, Humphreys does not give us a full explanation of causation, but theories of causation do actually vary within the scientific and philosophical communities. Further, his explanations involve incompleteness, but this too is appropriate, since our scientific knowledge is in fact incomplete.

Second, our model ought to involve causation essentially. This is a requirement satisfied by several of the models we have looked at, so Humphreys' model is not unique in this respect, but his account does clearly rest on causation. The elements of the sets Φ and Ψ must be causes.

Third, the model ought to allow for both deterministic and indeterministic explanations. This requirement is satisfied, since Humphreys allows the causes in question to be deterministic or indeterministic. This might raise some questions about whether there could be indeterministic causes, but the question is not at all harmful to the theory, since if there are no indeterministic causes, then one ought only to cite *deterministic* causes in explanations. This is not a problem, since the model works either way.

Fourth, the model should avoid citing natural laws unless it can give a good non-circular account of them. Humphreys satisfies the first part of this disjunction by leaving laws out of his model altogether. In fact, not only are laws not explicitly mentioned in the canonical form, it would be unlikely for a law to ever count as a member of Φ or Ψ, since laws merely describe regularities, and do not themselves *cause* anything.

Fifth, the model should avoid having any sort of requirement of maximal specificity (RMS). As discussed, failure to meet this requirement results in explanations that are either false or so overly detailed that they describe regularities of *only* the single event in question, so that most or all explanations are deterministic (and somewhat circular). Well, aleatory explanations are not required to be maximally specific, and, indeed, are almost always incomplete in some way. The requirement that the causes be properly sorted as contributing and counteracting is nowhere near as limiting or problematic as RMS (as discussed earlier in this chapter).

Sixth, our explanations should be objectively true (that is, not relativized to any epistemic situation) and empirically verifiable (or falsifiable). Certainly, aleatory explanations are objectively true. Whether or not a particular thing is a cause of another is not a matter of a knowledge-situation or a state of information. Nor is it a matter of any person's (or society's) opinions or beliefs. Presumably, whatever is meant by 'causes', they are real things "out there" in the world. Further, if causes are indeed "out there," they can probably be discovered and understood (reasonably well) through empirical observation and experimentation. As such, it would seem that Humphreys' account meets both parts of this requirement.

Finally, it is important that any explanatory account be essentially non-inferential. In the next chapter I will discuss the reasons for this in some detail, but for the moment, I will simply say that telling me something was *likely* to happen or *could be expected* to happen (as the D-N and I-S models do with regard to explananda, and as the D-N-P model does with regard to the *probability* of explananda) does not get me any closer to an understanding of *why* it happened. Humphreys clearly avoids this pitfall, eschewing not only inference itself, but also the assignment of probability values that would make such an inference possible.

I conclude, then, that the aleatory model of scientific explanations is superior to the other models discussed, as it answers to their problems without creating any undue difficulties in the process. If we further assume that these models are indicative of the sorts of models of scientific explanation that exist within the literature, then it would not be unreasonable to imagine that Humphreys theory is *the* best (or at least one of the best) models of explanation that there are. Whether or not this is true, I think that the aleatory model does work very well to describe and determine the explanatory practices of science, even though there are some minor modifications that must be made with regards to the *quality* of aleatory explanations (which I will propose in my last chapter).

My next two chapters discuss some of the reasons for and implications of this sort of explanatory model. In particular, I will discuss the motivation behind the non-inferential account in the next chapter, and I will discuss the merits and shortcomings of contrastive and non-contrastive explanations in the following chapter.

CHAPTER V

EXPLANATION VS. INFERENCE

The aleatory model of explanations is essentially non-inferential. That is, just knowing the explanation given of the explanandum, one could not infer (inductively or deductively) that the explanandum is in fact the case. Of course, it is always assumed that only truths can be explained, so in a sense, one could infer the truth of the explanandum *from the existence of* the explanation, but what is important here is that one could not infer it *from the facts contained* in the explanation. So given the canonical form "A because Φ, despite Ψ," one could know all the members of Φ and Ψ and still be incapable of inferring A. There are three reasons for this. First, when it comes to indeterminate causation, which is allowed on Humphreys' model, one is never *sure* that the effect will occur along with the cause, and in some cases, it is in fact very unlikely that the effect will occur along with the cause. Remember, even very unlikely events can be explained on the aleatory model. Second, and this applies to deterministic and indeterministic causation, explanations are pretty much always incomplete, so there could be some contributing causes involved that are not mentioned, and hence the inference could be blocked. Most importantly, however, for the purposes of this chapter, aleatory explanations are not inferential *in their format*. The explanandum is not presented as the conclusion of an argument of any sort.

The idea that explanations are a sort of inference, however, is fundamental to several other models, and is a commonly held idea within the philosophy of science.

Hempel, exemplifying the logical empiricist tradition in this matter, was a staunch advocate of this view, as we can see by the format of the D-N and I-S models that he presented us with. In this chapter I will discuss some of the reasons why explanations should *not* be thought of as inferences and why, therefore, Humphreys' model is not only an acceptable one (with regard to this matter), but superior to those inferential models mentioned earlier. It will also turn out that some of the reasons we have for thinking that explanations should not be inferential can be turned around and used as reasons why explanations *should* be causal (which Humphreys' model quite clearly is), so I will dedicate some time to this matter at the end of the chapter.

Jeffrey on Explanation and Inference

In "Statistical Explanation vs. Statistical Inference,"[38] Richard Jeffrey vigorously attacks the notion that statistical scientific explanations are arguments. He does so by pointing out a number of ways in which explanations and arguments seem, intuitively, to differ from one another. Before going into Jeffrey's arguments on this matter, it is important to point out one way in which his analysis of explanation differs from the analyses of many other philosophers. Where most philosophers (as I have previously mentioned) consider an explanation to be an answer to a 'why?' question, Jeffrey speaks of two sorts of explanation-seeking questions: 'why?' and 'how?'. He allows that the answer to either question might be considered an explanation, and that a *complete*

[38] Richard Jeffrey. "Statistical Explanation vs. Statistical Inference."

explanatory account should answer both questions[39]. The 'how?' question is answered in terms of whatever process it was that brought about the event in question. This will be some sort of stochastic (i.e. causal) mechanism, and could be either deterministic or indeterministic, provided only that it leads to the explanandum. The answer to the 'why?' question in statistical, or probabilistic, situations is always 'by chance', or, more interestingly, 'there is no damned reason, it was pure chance'. Although he is not explicit about this, I think it may be safely inferred that the answer to the 'why?' question in deterministic situations will always be something like 'because the stochastic process (from the answer to the 'how?' question) necessitated (determined) it'.[40] It is not entirely clear to that causal information always (or even usually) answers 'how?' rather than 'why?' questions, nor is it clear that an adequate explanation must answer to each of these questions. In fact, it seems to me that with regard to scientific explanations in particular, the 'how?' is usually included within the 'why?', and that one has simply not explained *why* an event occurred if one has made no reference to a cause or a stochastic process. However, the acceptance of Jeffrey's view of explanation-seeking questions is not crucial to one's acceptance of his arguments here. It is just important to be aware of it.

The first problem that Jeffrey brings up is hard to give a good name or description to. It arises in cases where an argument renders its conclusion probable, or

[39] Already we see a fundamental difference between Humphreys and Jeffrey. Humphreys sees no reason why there should ever be a 'complete' explanation in science, but Jeffrey seems to think it to be a realistic and desirable state of affairs.
[40] "In general, where the inference meets certain conditions, one of which is that a causal law appear among the premises, deductively grounded prediction will double as explanation. In these cases, knowledge *that* is imparted in such a way as to provide knowledge *why* as well." Ibid. 107

even certain, but fails to explain that conclusion because of the ways in which the premises relate to one another or to the conclusion itself. To use Jeffrey's example:

> The glass is falling.
> Whenever the glass falls the weather turns bad.
> ∴ The weather will turn bad.[41]

The above argument deals with true (or at least true for the sake of argument - obviously it doesn't really rain *every* time the glass falls) empirically verifiable premises, one of which is a general law. The conclusion is validly deduced, and both premises were required for the deduction (all of which is important since Jeffrey's main argument is against D-N and I-S models of explanation). However, we do not have an adequate explanation of the explanandum. The falling of the glass does not in any way bring about the change in the weather. The premises would, intuitively, never properly answer a 'why?' or 'how?' question about the weather. Another example is Bromberger's flagpole example.[42] We know that a certain flagpole is x feet tall, and that it casts a shadow that is y feet long. We also know the position of the sun at the time in question and using that fact, in conjunction with the law of rectilinear propagation of light, we can deduce the length of the shadow (y) from the height of the flagpole (x). This would, on Hempel's theories at least, count as a valid explanation. What is odd, though, is that we can turn the explanation around and explain the height of the flagpole in terms of its shadow (along with the relevant facts about the sun and the nature of light). This seems

---

[41] Ibid. 106
[42] Wesley Salmon. *Four Decades of Scientific Explanation.*

counterintuitive.[43] How can a shadow tell us *why* (that is *explain* why) the corresponding has the height it does? There seems to be a definite gap between explanation and inference in both these cases. Jeffrey is willing to admit, however, that arguments of the following form could serve as explanations[44]:

> The glass is falling.
> Whenever the glass is falling the atmospheric pressure is falling.
> Whenever the atmospheric pressure is falling the weather turns bad.
> ∴ The weather will turn bad.[45]

He thinks this is a fine example of explanation because it really lets us know *why* the event in question occurs. Here we have an explanation that refers to the stochastic process whereby the weather turns bad, and we therefore have a good understanding of that event. The first two examples shed light on the idea that inference (deductive or otherwise) is not a sufficient condition for explanation, even when all the proper Hempelian restrictions are taken into account, but the third leads us to wonder whether is might at least be a necessary condition.

It is in response to this notion that Jeffrey gets into probabilistic explanations. He uses, as an example of a probabilistic (or *chance*) event, a series of coin tosses. If we

---

[43] It is interesting to note that Hempel himself had no problem at all with such explanations. He uses the example of an explanation of the period of a pendulum in terms of its length being turned into an explanation of its length in terms of its period. While he admits that the latter explanation seems odd to our pre-philosophical and pre-scientific intuitions, he thinks that once we come to terms with the fact that *you cannot change one without changing the other*, the oddness goes away.

[44] Though I would disagree, along with Humphreys, that it is an *explanation*, since the relevant causal information does not fit the canonical form. I will, however, follow Jeffrey's arguments through without worrying too much about the precise format of his explanatory accounts. He is not proposing an alternate model of explanation in this work, but is rather working within that Deductive-Nomological framework in order to attack it.

[45] "Statistical Explanation vs. Statistical Inference." 106

toss a coin 10 times, with a probability of .5 of heads on each toss (where the tosses are, of course, independent of one another[46]), we can deduce ahead of time what the probability of any given combination of results will be. For instance, the probability of getting zero heads after ten trials is 1/1024. The probability, then, of getting at least one head will be 1023/1024. Jeffrey generalizes the probability calculations for "at least one head" in the following manner:

If $i_1, \ldots, i_k$ are all distinct, then $P(Hi_1 \& \ldots \& Hi_k) = 1/2^k$
=======================================[1 -2^{-n}]
$H_1 \vee H_2 \vee \ldots \vee H_n$

Where H_i = "toss number *i* comes up heads'

Now Hempel might regard the above argument form as an explanation of the occurrence of one head, provided that n is a large enough number. For instance, if we flipped the coin ten times the argument above would have a strength of 1023/1024, which would almost certainly be a sufficiently high probability for Hempel. If we only flipped the coin once, however, the strength of the argument is only 1/2, which is insufficient for Hempelian explanation. But, says Jeffrey, there is a confusion involved here between the probability of the explanandum and the *strength* of the argument. Jeffrey sees the above argument as a deductive argument conferring the probability of 1 - 2^{-n} on its conclusion. It does not, however, answer any questions about *how* or *why* the explanandum-event occurred.

---

[46] So $\forall xy[\ P(H_x \mid H_y) = P(H_x \mid \neg H_y)\ ]$, or , the probability of heads on any one trial given heads on any other trial is the same as the probability of heads on that first trial given not-heads on the second. That is,

> Perhaps it is part of the explanation to point out (as in the brackets above) that the probability of the explained phenomenon is…1 - $(1/2^n)$, but the strength of the explanation would be no greater with n = 10 than with n=1! To say, in the case n=1, that the statistical probability is 1/2, is as much of an explanation as one can give.[47]

Jeffrey's basic argument for why the occurrence of at least one head on a single toss is just as explicable as the occurrence of at least one head on ten tosses is simply that in either case we understand the statistical and the stochastic processes involved. Furthermore, because of the nature of the stochastic process in question, which is a chance process, the argument given above does not in any way add to our understanding. It is really deductive in a sort of trivial way (where the conclusion is quite clearly contained in the premise), so it can hardly be said to answer any 'why?' or 'how?' questions for us. As Jeffrey puts it:

> [F]or n=10 it is possible, although highly unlikely, that there will be ten tails, and if this happens we shall know all there is to know about the why of it and the how, when we know that the process which yielded the ten tails is a random one and when we know the probabilistic law governing the process. The knowledge that the process was random answers the question, 'Why?' - the answer is, 'By chance'. Knowledge of the probabilistic law governing the process answers the question 'How' - that answer is, 'Improbably, as a product of such-and-such stochastic process. Note that knowledge of the probabilistic law of the process makes it at most a matter of calculation to find that the statistical probability of the phenomenon - ten tails - is 1/1024.[48]

Hence, it would seem that arguments are not only insufficient for explanations (on Jeffrey's account), they are also unnecessary. Once we understand that the event in question occurred (as Jeffrey puts it) 'by chance', and 'Improbably, as the result of such-

---

the outcome of one event doesn't change the probabilities of any of the others.

[47] Ibid. 108

[48] Ibid. 109

and-such stochastic process', we really have all the answers we need as far as explanation goes.

Having established that explanations are not arguments, Jeffrey sees the need to provide an error thesis to explain why we are so convinced by the idea that explanations *are* essentially inferential. The answer is simply that it is usually possible, when an event occurs, to give an argument that confers at least reasonably high probability on it. As for why this is the case, there is a tautological (or apparently tautological) explanation: the usual usually occurs. Whenever an event occurs, more often than not, it will be probable; that is just what it means for something to be probable (at least on most accounts. Remember that Humphreys does not think it is appropriate to speak of the probability of a chance event). In Jeffrey's words:

> The situation [is not] changed when (as usual) the probable happens. Because such cases are usual, we can usually give a statistical inference of strength 1/2 or more then we can give a statistical explanation; and this, I take it, is why it is easy to mistake the inference for an explanation.[49]

Although I agree with Jeffrey's conclusion that the view of explanations as arguments is in error, I do not think this argument adequately explains the error.

First of all, it is not at all clear to me that most events that occur are probable or "usual." Such a thesis, I think, assumes a very odd view of events, both probable and otherwise. To demonstrate this, imagine the roll of a single, six-sided die. I take it that the roll of a die is a very good and ordinary example of a chance event[50], and that the

---

[49] Ibid. 109

[50] It is certainly used often enough in introductory texts on probability.

most natural way of stating the possible outcomes is the set {1,2,3,4,5,6}. The result of the roll will be one of these six outcomes, with an equal chance of each. But of course, that means *none* of the possible outcomes are at all probable. The only way to get a probable event out of the roll is to subsume certain outcomes under broader definitions in a rather counterintuitive way, like {1, not-1}. In this case, not-1 is far more probable than 1, but we have lost a certain amount of depth in our analysis of the probabilities. To be sure, there are cases where {1, not-1} are the only results that matter to me (like if my life, for some reason, depended on a particular die coming up 1), but we cannot always be satisfied by such options. There are, of course, many other ways of setting up the probabilities ({less than 3.5, greater than 3.5}, {odd, even}, {factor of 6, not a factor of 6}, {1,2,3,4, greater than 4}, and so on) some of which will render the actual result probable, others will render it improbable.

If even a reasonable number of the probabilistic events in the world have more than two logically distinct ways of describing their possible outcomes[51], then whether or not the majority of chance events are probable would seem to be a matter of context and interest. It is not up to the logician or the statistician to determine ahead of time what outcomes are most important, and so it is not possible to say *a priori* what events are

[51] What sort of case qualifies for this exception? Well, take the case of a coin weighted such that P(H) = .6 and P(T) = .4. There are only two logically distinct ways of describing the outcomes of this trial that are consistent and that exhaust all the possibilities: {H,T} and {H v T}. {H v T} is, of course a non-probabilistic description of the outcome, but it does accurately describe the outcome of the trial. I will accept that any situation in which there are only two mutually exclusive and jointly exhaustive outcomes with different probabilities (like the aforementioned weighted coin) is a situation in which there is a probable or "usual" event, but I do not think that such cases are numerous enough to imply that most events are probable.

most probable or "usual," or indeed, whether or not there are *any* probable events[52]. Jeffrey may still have a good point here, if it is indeed the case that we ordinarily subsume events like the outcome of a die, the flipping of the penny, or the fluctuations of the market under broader (i.e. more probable) rather than narrower (i.e. less probable) descriptions, but he would be leaving open the question of *why* we behave in this way[53], which is, I think, an important part of any error thesis.

My second point of contention with Jeffrey deals with the logical structure of his argument. He has said that most events are *usual*, and that when an event is usual we can *usually* give an argument for it. Even if we were to grant these premises, the conclusion that we can give an argument for most events does not follow from them. Imagine that "most events are usual" is true because the usual occurs 60% of the time (that is, with .6 likelihood[54], a given event that occurs will be usual). This would satisfy Jeffrey's requirement that the probability be 1/2 or more. Further, imagine that "we can usually give a statistical inference of usual events" is true because we can give a statistical inference of usual events 60% of the time (that is, the probability we can give an inference of a usual event is .6). From these two premises the conclusion that follows is that we can give inferences for only 36% of all events.

---

[52] Except, of course, to the extent that he or she could show there to be cases like that of the weighted penny, but this would require research beyond what the logician does *qua* logician.
[53] Which, I take it, must be answered psychologically. That is, it must be answered by a psychological *explanation.*
[54] This is, I admit, a very odd way of speaking, since I am using the word "most" (i.e. >50%) along with a statistical value of .6 (which, of course, is >50%). This may seem to be trivially true, but I do not think that it is. See my argument in the preceding paragraphs.

P(Ux) = .6
P(Ix | Ux) = .6
P(Ux & Ix) = .36

x is an arbitrary event
Ux = 'x is usual'
Ix = 'x can be inferred'

So it is clear that the argument, as it is set up, is inadequate to explain the error that is committed by those who think that inference is a necessary component of explanation. How then can we solve the problem? Certainly if one were willing to admit that the *vast* majority of events are "usual," or that a vast majority of "usual" events are inferable, Jeffrey's argument would be valid. But for reasons that I have explained above, I find such an approach highly questionable, and, at any rate, very likely to be unsuccessful.

I think the more reasonable approach is to ask ourselves whether we do ordinarily believe that explanation requires argumentation. I think that a moment's reflection on either our first experience with the D-N or I-S models or on the commonplace acts of explanation that we both witness and perform on a daily basis will reveal just how little the notions of argumentation and inference *have* affected our notion of explanation. We consider environment, diet and genetics to be good explanations of certain cancers, despite the fact that these factors do not confer a high probability on the cancer itself. Even smoking, the American Lung Association's number one "no-no," raises the probability of getting lung cancer by considerably less than 50%, which means that one could never provide a convincing argument that smoking will be followed by lung cancer, and yet we regularly explain lung cancer in terms of smoking habits. These preceding examples show our lack of interest in inference and predictability in some commonly understood *scientific* explanations, which is, of course, important to my thesis. But for the purposes of analyzing our most naive intuitions, let us take a far more banal example. Suppose I walk up to a bank of three elevators, and press the button. The elevator in the middle soon opens up and I get inside. A perfectly good explanation of why the middle elevator opened is *that I pressed the button.* This is despite the fact that there was only a one in three chance of that particular elevator opening.

On the other side of things, we very rarely consider arguments *for* an event to be at all adequate in *explaining* the event. If I ask, "why does the television come on when I press the button on the remote?" and get the response, "These are a late-model television and remote. The television is plugged in and the remote has fresh batteries. In such circumstances, the probability that the remote will activate the television is greater than 99%. That means that the event in question was very likely, and so it occurred." I would respond with, "yes, that's all well and good, but I want to know *why* the television turns on when I press the button on the remote?"[55] Such arguments are often completely unconvincing as explanations, despite the high (sometimes even deductive) likelihood that they confer on their conclusions.

The purpose of going through our intuitions about all these cases (Jeffrey's, and my own) is to show that we really don't find Hempelian argument-style explanations all that convincing in the first place. No matter how committed a Hempelian one is, it is likely that in ordinary practice non-argumentative explanations like the ones above seem

[55] Jeffrey, of course, would say I'm looking for the answer to the 'how?' question here, but whether it's 'why?' or 'how?', the argument clearly doesn't provide an answer.

pretty good (and conversely, some of the argumentative ones seem pretty bad).[56] It is odd that Jeffrey thinks we are as attached to inference as he does, since most of his article is designed to reveal our intuitions on the subject, and they all come out negative. So, to restate: the error is not our conviction that good explanations must be inferences, it is the belief that we ever had such a conviction in the first place.

Salmon on Explanation and Inference

Wesley Salmon, in *Four Decades of Scientific Explanation,* building on Jeffrey's work, provides us with two more reasons to think that explanations are not arguments. First, there is a problem with irrelevancies in explanation that does not come up in argumentation. If I try to explain Jones's failure to become pregnant by saying that he is a male who faithfully consumes birth control pills, my explanation is a bad one. His consumption of birth control pills is, of course, totally irrelevant to his not becoming pregnant, and this hurts the explanation. If I just say "Jones is male," I have a good explanation, but when I add the birth control pill irrelevancy, I get a bad explanation. To consider a less frivolous case, imagine that a drug is administered to a person suffering from a common cold. After a few days their cold symptoms clear up completely. It would be very unwise (and, in fact scientifically irresponsible) of us to explain the recovery by mentioning the drug. After all, most common colds clear up after a few days with or without treatment, so without any (extensive) research into the relative

[56] One might argue that they only seem good *when we are not thinking philosophically,* and that our everyday discourse can be separated from our philosophical views. To quote David Lewis, "An analysis that imputes widespread error is *prima facie* implausible. Moreover, it is dishonest to accept it, if you yourself persist in the "error" when you leave the philosophy room." (David Lewis. "Causation," 175)

recovery speeds of those administered the drug versus a control group, it would be unwise to consider the drug relevant, and mention of the drug therefore lacks explanatory force. But this is very different from any sort of valid deductive inference[57], where irrelevancies are harmless. Consider the following:

> Either Jones is not pregnant or Mary is not pregnant
> Mary is pregnant
> <u>Jones has faithfully consumed birth control</u>
> Jones is not pregnant

> All people with common colds feel better after a few days
> Jones had a common cold a few days ago
> <u>Jones has taken the new miracle drug</u>
> Jones feels better

These are both perfectly valid sequents. Neither one, however, could work as an explanation because they contain irrelevancies in the third premise. The argument Salmon gives here is intended to build on the points Jeffrey made in the barometer example. Not only can an argument not explain when the premises relate in improper ways to the conclusion (as Jeffrey notes), but they can also fail as explanations when there exists a premise that does not relate *at all* to the conclusion (That is, the conclusion is deducible without the premise). Hempel is, of course aware of this problem, hence the restrictions he puts on D-N explanation. Even with that restriction in place, however, Salmon's point still holds that there is a definite asymmetry between explanations and arguments.

---

[57] This applies even to relevance logic in some respects. While P ∴ ¬(Q & ¬Q) is not a valid sequent in relevance logic, P & Q ∴ ¬(Q & ¬Q) is, despite the irrelevancy of the P in the premise. An explanation

The other problem Salmon presents us with is that of the differing temporal restrictions of explanations and arguments. An argument for the occurrence of event e could, it would seem, contain vital premises regarding the nature of things *after* e has occurred. For instance, the occurrence of a solar eclipse can be inferred from the positions of the earth, sun, and moon prior to the eclipse, along with the relevant laws of planetary motion. An inference of this sort would, at least on some accounts, count as an explanation of the eclipse. However, the occurrence of an eclipse could be just as easily inferred from the positions of the sun, earth, and moon *after* the eclipse. We would not want to say, however, that the latter inference could count as an explanation. This may be a somewhat controversial point, but it seems to have a great deal of intuitive appeal, if only from the ways we ordinarily use language, and the way we regard explanations as answers to "why?" questions. There are many different diseases that we say can explain a person's death, but there is never a case in which we say that a person's death (no matter how distinctive it is) *explains* their having had a particular disease (or having had a disease in general), even if we can infer their having had the disease from the fact of their death. Sherlock Holmes may infer that a particular person committed a murder from the way they behave after the fact, and we might say that their having committed the murder explains their behavior, but neither Holmes nor we would ever say that their behavior after the fact *explains* their having committed the murder. After all, the statement "because he acted in a certain way afterwards" does not answer "why did the

---

would seem to fail if even a single conjunct of a premise were irrelevant. Consider: Jones is male & Jones takes birth control.

murder occur?,"[58] and "because he died" could never explain "why did he get sick?" So an explanation, unlike an argumentative inference, must make some appeal to things that occur *prior to*, or at least *at the same time as*[59] the explanandum.

It is interesting to point out here that David Lewis (whose view of causal explanation I have mentioned) does not think that explanations can never cite events that occur after the explanandum. His reason for this is that he believes time travel to be at least logically possible, and that causation therefore has the potential to go "backward." This exception is not a devastating one with regard to Salmon's experiment, since the examples involved (the eclipse, for instance) do not in fact involve time travel. It is, however, a very interesting thing to keep in mind, if only because of the relativistic concept of time that modern physics presents us with[60] — one would not want to make any undue assumptions about the physical nature of the universe a part of one's account of scientific explanation.

So where *do* our intuitive beliefs about explanations lie? Well, Jeffrey and Salmon will both claim that what is really important to any explanation is that it gives us some understanding of the *causes* (or the stochastic process) involved in bringing about the explanandum. Both writers find this to be a rather intuitively appealing way of understanding explanation and don't provide much in the way of explicit argument for it

---

[58] Or perhaps "why did he commit the murder?"

[59] Whether the explanans and the explanandum can occur simultaneously is not entirely clear to me, but as long as it is understood that the explanans cannot come *after* the explanandum, we have the necessary asymmetry between explanations and arguments.

[60] General relativity tells us that neither time nor space is an absolute, but that they are both relative to speed. In other words the space and time in which an event occurs depends strictly on its relative speed (relative, that is, to the speed of light in a vacuum). It is not the purpose of this thesis to sort out the philosophical implications of the theory of relativity (good thing, too, because otherwise I would be failing

(with the exception, perhaps, of Jeffrey in the coin-tossing case), but I think it would be appropriate to examine how the asymmetries that they noticed between explanation and inference can be used as the premises for an argument with causal explanation as its conclusion. The argument will, of course, be an abductive[61] one, so there may well be arguments of similar strength that have opposite conclusions, but I will be happy if I can at least show how the notion of causal explanations can tie together our intuitions regarding the examples above.

First, we have Jeffrey's claim that an explanation of *why* something happens must tell us something about *how* it happens. The barometer, taken in isolation, simply cannot tell us why it rains. When we add information about the meteorological situation just prior to rain, however, we can really *understand* the situation. Atmospheric pressure is connected to rain in a way that barometers just aren't. It is not easy to say exactly why, but perhaps it is just the fact that the clouds are way up in the sky and the barometer is down on the ground. It would certainly be difficult to explain the problem in purely formal terms. Salmon's point about the damaging effect of irrelevancies on explanations brings up a similar issue, and, like in Jeffrey's example, it is hard to say what exactly is so different about the arguments with irrelevant premises and those with only relevant premises.[62] If we imagine that causation is the crucial aspect of explanation, however,

---

miserably), but this does cast some doubt on the correctness of our naïve intuitions about the temporal direction of explanation.

[61] Abduction is, very simply stated, the logic of discovery, or of hypothetical inference. So the premises will not make the conclusion certain, likely, or even possible, but rather the conclusion, if true, would render the premises understandable..

[62] Interestingly, Hempel has no problem explaining this case away. He would simply say that the irrelevant premises are not required in the deduction of the conclusion. Salmon and Jeffrey do not have this line of reasoning open, however, since they do not believe explanations are arguments.

we can easily make sense of both these examples. Why does a barometer not explain the weather? Because it doesn't *cause* the weather. Why *does* atmospheric pressure explain the weather? Because it *causes* the weather. Similarly for Salmon, the irrelevancy problem would be solved if one were to claim that all and only causes properly explain. Since a cause cannot fail to be relevant to its effect, we can now argue for why the irrelevancies are harmful. Since the consumption of birth control pills has no causal effect on a man's ability to become pregnant, it should not be included in the explanation.

Jeffrey's next point deals with low-probability explanation. We can explain even those events that are statistically very unlikely, provided that we understand the process through with they occurred. Here Jeffrey himself is quite explicit, when he says that the ten-tails result occurs "Improbably, as the result of such-and-such stochastic process." Jeffrey tells us that what is missing from the purely statistical inference is the *causal* part of the explanation. This hypothesis (that causes are crucial to explanation) does seem to explain how it is that we *understand* even the highly improbable. When we get ten heads in a row, we are surprised, but we are not confused.

Finally, we have Salmon's temporal priority requirement on explanations. The position of the sun, moon and earth *before* and eclipse can explain the occurrence of the eclipse, and those same positions *after* the eclipse cannot, despite the fact that the eclipse can be deduced from either situation. Here again, the causal hypothesis offers us a clear reason for our intuitions here. Earlier events can cause later events, but later events cannot cause earlier events (this is not, of course, uncontroversial. Refer to my earlier comments about David Lewis's view of causation), so the latter argument lacks the explanatory force of the former.

The examples that I have examined here suggest that an explanation must 1) be temporally prior to the explanandum, 2) be entirely relevant to the explanandum (that is, it must contain no irrelevancies), 3) tell us something about the process through which the explanandum occurred (that is, it must answer a "how?" question), and 4) it must give us an *understanding* of the explanandum. It seems quite clear to me that these four conditions lead us to the conclusion that causation is a crucial part (if not *the* crucial part) of any explanation.

Humphrey's account then, has the advantage of being both essentially non-inferential and essentially causal. This allows it to avoid many of the pitfalls suffered by models such as Hempel's and it follows more closely our intuitions on the matter of explanation.

CHAPTER VI

CONTRASTIVE EXPLANATION AND THE SYMMETRY THESIS

In this chapter I will be discussing two issues as they relate to Humphreys' model of aleatory explanations. First, I will explain why contrastive explanations are not necessary in scientific discourse. I will do this primarily through a discussion of Bas van Fraassen's model of explanation. This model was not mentioned earlier, primarily because it is not a model of *scientific* explanations in the sense described in the introduction, but rather of explanations *in* general[63], but also because the detailed examination that I will be giving it in this chapter, and my desire not to repeat myself. The second issue is that of Hempel's "symmetry thesis." This is the idea that explanations must be capable of serving as predictions and vice versa. As aleatory explanations will rarely meet this criterion, I must show why it is not as essential as Hempel thought (or perhaps why it is not essential at all). These two requirements (that they be contrastive and that they adhere to the symmetry thesis) on explanatory models are distinct, but I believe they have similar motivations. I will discuss these motivations briefly at the end of the chapter and explain why they are inappropriate.

---

[63] Because of this, when discussing Bas van Fraassen's theory of explanation, I will not restrict myself to scientific discourse, as he does not. It should be noted, however, that his model is intended to cover scientific explanations, as well as other types.

Contrastive vs. Non-Contrastive Explanation

The appropriate answers to a given question in English are in part determined by the way one asks the question and by its context of utterance. Even within the relatively rigid language of scientific discourse, two or more 'why?' questions that are syntactically identical (i.e. they have the same words and punctuation in the same order) can nevertheless be asking for very different information. Consider the three sentences:

a. Why did *Ms. R.* kill Mr. R.?
b. Why did Ms. R. *kill* Mr. R.?
c. Why did Ms. R. kill *Mr. R.?*

Each of these questions are syntactically the same, but the emphasis put on the parts of each makes it clear that the same answer will not satisfy all three questions. An acceptable answer to (a) might be "because when we decided that Mr. R. had to be killed, straws were drawn and she drew the short straw," but this clearly does not provide us with an answer to either of the other two questions. Similarly, "because she could make more money off the death of Mr. R. than she could off of the death of Mr. S." would be a fine answer to (c), but fails as a response to the other two. For a more scientific example, consider the case raised by Humphreys of the room that is both air-conditioned and heated. When we ask, "why did the temperature rise by 5˚C?" we might be wondering why it rose at all or why it rose by exactly 5˚C.

Bas van Fraassen has developed an entire explanatory system for handling 'why?' questions that are syntactically identical, but that have different "acceptable" answers.[64] He starts with the notion that any explanatory question has an appropriate "contrast class" of statements that are of the same logical form as the statement to be explained. Which statements are to be included as members of the contrast class is a matter to be determined by how the question was asked, or what sort of information it is seeking. For instance, a contrast class for (a) might be {'Ms. R killed Mr. R.', 'Ms. S killed Mr. R.', 'Mr. S killed Mr. R.', 'Mr. R. killed himself.'}, and the contrast class for (b) might be {'Ms. R killed Mr. R', 'Ms. R hugged Mr. R', 'Ms. R yelled at Mr. R', 'Ms. R bought ice cream for Mr. R'}.

According to van Fraassen, all but one of the members of a given contrast class will be false, and an acceptable explanation will pick out the true one (or the "topic") according to a certain "relevance relation." The notion of a relevance relation is not simple to spell out, but it is basically the *manner* in which an explanation picks out the topic from the rest of the contrast class. The answer to any "why?" question must give a reason why the topic is true (rather than the other members of the contrast class), and the relevance relation puts a restriction on *what sorts of reasons will* count. To give it such a definition might make the first sentence of this paragraph seem trivially true (which it might be), but consideration of some examples might help us get a grasp on what we are talking about. In all three of the cases above we are most likely seeking some sort of

[64] Bas van Fraassen. *The Scientific Image.*

motive of Ms. R's, so our explanation will provide us with a motive. In the case of John's failure to contract malaria (from the chapter on explanatory models), we are likely seeking some physiological facts about John, so our explanation will provide us with some physiological facts. So the relevance relation tells us something about what the explanation will look like. Van Fraassen's total account is set up as follow[65]s:

[T]he why-question Q expressed by an interrogative in a given context will be determined by three factors:

The topic P_k
The contrast class $X = \{P_1, \ldots, P_k, \ldots\}$
The relevance relation R

and, in a preliminary way, we may identify the abstract why-question with the ordered triple consisting of these three:

$Q = \langle P_k, X, R \rangle$

A proposition A is called relevant to Q exactly if A bears relation R to the couple $\langle P_k, X \rangle$.

We must now define what are the direct answers to this question. As a beginning let us inspect the form of words that will express such an answer

(*) P_k in contrast to (the rest of) X because A

What is claimed in answer (*)? First of all, that P_k is true. Secondly, (*) claims that the other members of the contrast-class are not true…Thirdly, (*) says that A is true. And finally, there is that word 'because': (*) claims that A is a reason.[66]

Perhaps I owe it to the reader to spell out in more detail exactly what all of this means, but in the interest of time, I will not be doing this. The important thing about van

---

[65] I quote Van Fraassen at length here since his account is sufficiently brief and self-explanatory, and I do not think I could do much better.
[66] Ibid. 9

Fraassen's account, for the purposes of the present discussion, is that it is *essentially contrastive* in nature. A contrastive explanation is given in response to a question of the form "why the explanandum rather than this or that other thing?" Such a question is an example of a "why?" question, but it is clearly very specialized. It seeks far more information that an ordinary "why?" question and is, as a result, far more difficult to answer. Van Fraassen's account is *essentially* contrastive because this contrast is built into the very format of his explanatory model. He takes it as a very basic fact that explanation involves contrasting the actual fact with the relevant alternatives.

Contrastive "why?" questions present a major problem for any theory of explanation (like Humphreys') that allows for chancy (i.e. non-deductive) explanations, particularly those theories that allow for the improbable (or probable but not *very* probable) to be explained or make explicit reference to probabilistic causation. When low-probability events are in question, there can be no answer that explains "why…rather than…", since the other things in question (the other members of the contrast class) are jointly more likely to occur than the explanandum. If there were any such reason (why the low probability explanandum occurred, rather than the other alternatives), then the probabilities would have been misassigned from the start, and the explanandum was more likely that we thought. Further, Under a theory of probabilistic (chancy) causation, such as one might see applied to Humphreys' model, it doesn't make sense to ask why one event occurred rather than another if they could both have come about via the same causal chain. If A and B are two mutually exclusive possible effects of the same causal chain, and A occurs rather than B, then there is no reason why A occurred rather than B. Since the explanation cites only the causes, and there is no difference between the causes of A and the causes of B, their explanatory account would be the same whichever happened. If the probability of A is high, we might say it was to be expected, but even this is not a *reason* that it happened. To quote Jeffrey, "there's no damned reason – it was pure chance"[67]

Yet we obviously do ask contrastive "why?" questions. Does a causal theory of explanation that allows for the explanation of improbable events, such as Humphreys', make all such questions unanswerable? The answer to this question depends fundamentally on the nature of the universe, and cannot be determined a priori.

Let us look at Humphrey's aleatory model to see if it could be used to answer contrastive "why?" questions. The form of an explanation is "A because Φ, despite Ψ" where Φ is a non-empty set of contributing causes, Ψ is a set (possibly empty) of counteracting causes, and A is a sentence describing what is to be explained.[68] We define the notion of a contributing cause as (roughly) "something that makes the explanandum more likely" and a counteracting cause as (roughly) "something that makes the explanandum less likely."

Say I drop a raw egg off the roof of my apartment onto the concrete below, and it breaks. Further, let us say that such an event is completely deterministic, and that there was no probability of the egg's not breaking given its having been dropped. We could set up an aleatory explanation of this event as follows: "the egg broke because it was left

[67] "Statistical Explanation vs. Statistical Inference." 110.
[68] Paul Humphreys. " Aleatory Explanations," 227.

unsupported at a height of x feet, the Earth has a gravitational pull of y, and the shell of the egg has a disposition to crack whenever z amount of pressure is applied to it" This would be a true explanation. Since we are imagining that this event is *completely* deterministic, these causes must jointly confer a probability (likelihood) of 1 on the egg's breaking. I have not included any counteracting causes (wind resistance, gravitational pull of the moon, etc.) though a more complete account might (and could still be deterministic, as discussed above), but remember that the set Ψ can be the empty set.

I take it that in a case such as the one described, we have an adequate answer to any contrastive "why?" question regarding the explanandum. Questions like, "Why did the egg break rather than bounce off the concrete, break through the concrete, or sprout wings and fly away?" or, more simply "Why did the egg break rather than not break?" Since the contributing causes cited confer upon the explanandum a probability of one, we may safely infer that they also confer upon any event mutually exclusive with the explanandum a probability of zero. Hence, our aleatory explanation enables us to say, via two separate explanations, "The egg broke because of the causes mentioned (gravity, fragility, height, etc.), and it didn't fly away because of those same causes." So, it would appear, we have built up a contrastive explanation of the egg's breaking out of aleatory explanations. The egg broke rather than fly away because of the causes mentioned.

Of course, this is all well and good, but it assumes determinism, which would be a very odd thing for this model to assume indeed. While it is true that in our *example* determinism is assumed, it would be a grave mistake to think that the *model* assumes determinism. The model of aleatory explanation is neutral on the matter of determinism, so if there are deterministic events, it can explain them, and if there are indeterministic event, it can explain those too. The reason I used a deterministic case is that indeterministic events, as I explained above, *cannot* be given contrastive explanations. This is not just true of the aleatory model, but of van Fraassen's model as well. When all of the members of a contrast class are possible outcomes of a purely indeterministic cause, then there is *in principle* no reason why one occurs rather than the others. If we say that such a reason *must* be given in order to properly explain, then indeterministic events are all completely inexplicable, regardless of their probabilities. Now we can see why the answer I gave to "does Humphreys' causal theory of explanations make contrastive explanations impossible" was that it "depends on the nature of the universe." If there are indeed deterministic events, then Humphreys can offer contrastive explanations of them. If there are not any such deterministic events, then he cannot give contrastive explanations, but, I contend, neither could any other model, so Humphreys' is far from alone here.

But do we have good reason to imagine that it is important to be able to give contrastive explanations in science? There does seem to be a way of looking at many ordinary and scientific "why?" questions that makes many of them contrastive. We sometimes think that the question "why did the car skid off the road?" has an implied "rather than" clause of "rather than go straight." Certainly, if one takes explanation to be deductive (in D-N sense), this claim will have a certain appeal because of the frequent use of the *reductio ad absurdum* rule in deduction[69]. This rule assumes the opposite of the conclusion and derives a contradiction from it, thus proving the conclusion; which makes the opposite of the conclusion quite relevant to explanation. Also, in non-scientific teleological "why?" questions, the contrast class is usually very important. "Why did you pick *that* shirt to wear?," "why did you go straight rather than taking the exit ramp?," and "why are you drinking *American* beer?" are all seeking contrastive explanations. Some might imagine, as van Fraassen seems to, that this implied "rather than" clause is thus present in *all* of our "why?" questions, even when we are not aware of it.

But does the fact that we sometimes mean "why…rather than…" when we say "why?" imply that we *always* mean "why…rather than…?" when we ask "why?," or that we always *should*? It most certainly does not. The simple reason for denying it is that we sometimes ask "why?" and we sometimes ask "why…rather than…" Why would we have two different formulations of "why?" questions if they both meant the same thing? It is sometimes the case that the English language plays tricks on us, but we would need far more evidence than we have to think that that is the case with "why?" questions. I think it is entirely appropriate for a model of scientific explanation to be capable of handling both contrastive and non-contrastive explanations and "why?" questions. We should leave the determination of what sort of question is being asked to the scientists (or those asking questions of the scientists).

[69] In the previous chapter I argued that this is *not* an intuitive way of thinking about explanations, but,

The Symmetry Thesis

Carl Hempel, in *Aspects of Scientific Explanation*[70], presents the idea that an explanation and a prediction are essentially the same sort of thing, and that the only difference is a pragmatic one[71]. So, if sentences P and Q provide us with an explanation of R, then, if we had been aware of P and Q prior to the occurrence of R, we would have been able to predict R. Similarly, if we had predicted R on the basis of P and Q, then, should R in fact occur, we would have an adequate explanation of it (namely P and Q). This notion, because it posits a "symmetry" between explanation and prediction is often called "the symmetry thesis," and has become a popular criterion on the adequacy of explanations and predictions. Clearly, if it were an appropriate criterion, Humphreys' view of aleatory explanations would fail because of it. After all, with the allowance of incompleteness, not even a purely deterministic event could be predicted simply by its aleatory explanation, and in the indeterministic case, it would be even more difficult. I will examine here some of the common counterexamples in the literature to both parts of the symmetry thesis, and will add a few of my own. The conclusion will be that there is no good reason to accept the symmetry thesis, despite Hempel's claims to the contrary.

One of the main reasons that the symmetry thesis has enjoyed the popularity that it has is that many (including Hempel) believe it can properly distinguish between

---

nevertheless, there are those who hold this view, and they may, for that reason, believe in contrastive explanations.

[70] Later in his life, Hempel dropped the second half of the symmetry thesis (that predictions are essentially explanatory) because of some of the examples I will be presenting here. Hempel never, however, gave up on the idea that explanations are essentially predictive, and this is the aspect of the symmetry thesis that is most important to *this* thesis.

[71] Carl Hempel. "Aspects of Scientific Explanation."

science and pseudoscience. So an astrologer might correctly predict that a given person will "find true love" during a given time period. The symmetry thesis tells us that this is not a scientific prediction because it has no explanatory force. When asked *why* her prediction came true, the astrologer might respond that "Mercury was rising into Capricorn" or some such nonsense, but she will not be able to give an adequate explanation of the true-love-finding event according to *any* of the models explained above, and this lack of explanatory power is sufficient (according to the symmetry thesis) to make her prediction non-scientific.

The converse of the symmetry thesis can be used against such apparently well thought out explanations as those found in psychology. I don't wish to imply that psychology and astrology are on the same footing, but the excessively teleological nature of most psychological discourse precludes it from counting as a *science* within the language of this paper[72]. Peter Achinstein uses Freudian psychology as an example of just the sort of thing the symmetry thesis was meant to exclude. I give his example here because of the light I feel it sheds on the nature and purpose of the symmetry thesis:

> Freud explained slips of the tongue as the outcome of a compromise between an unconscious wish for something and a conscious wish that precludes it. He tells the story of the president of the German parliament who opened the meeting by saying "the meeting is now closed." This slip of the tongue, Freud explained, was the result of a conscious wish to open the meeting and an unconscious wish to avoid the meeting altogether. By using the expression "The meeting is now…" the president expressed some element of the conscious wish to open the meeting. But by replacing the word "open" with the word "closed," the slip contains a crucial element of the unconscious wish.
>
> Is this an adequate scientific explanation of why the president said what he did? Carl G. Hempel…claims it is not…Suppose before the slip of the tongue one had known

[72] See my introduction for the restrictions on what I am regarding as 'science' and 'scientific explanation'.

that the president had a conscious wish open the meeting and an unconscious wish to avoid it. Using Freud's theory could one have predicted that the president would say "The meeting is now closed"? No, because Freud's theory concerning slips of the tongue does not say that a slip of the tongue always or even usually occurs when one has conflicting wishes…Nor does the theory provide enough detail to infer what type of slips can occur…He suggests some explanatory ideas, but there needs to be a much more fully developed theory before Freud can adequately explain particular slips of the tongue.[73]

This is a case in which we can exclude a pseudo-scientific explanation on the basis of its non-predictive power. The intuitive motivation for this is captured in the question "how can a theory properly claim to *understand* an event if it can only ever do so *after* the fact?" It seems parallel to the man who says he knows everything and challenges anyone to find a true sentence he doesn't know. When presented with any example he merely responds, "Oh, yeah. I knew that."

However, despite the appeal of the symmetry thesis, I hold that neither side of it is in fact correct, and that there could be non-explanatory predictions and non-predictive explanations. I will first give some counterexamples to the idea that predictions are essentially explanatory and then devote a bit more time to some counterexamples and other replies to the idea that explanations are essentially predictive. This latter part of the symmetry thesis would, if true, be fatal to Humphreys' account of explanation, so it imperative that I determine whether or not it is true.

Prediction as Essentially Explanatory:

For many centuries before the birth of Newton, mariners and astronomers were

[73] Peter Achinstein. "The Symmetry Thesis."

aware of a correlation between the position of the moon and the tides. As we now know, thanks (at least in part) to Sir Isaac, there is a direct casual connection involved here--the gravity of the moon "pulling" the water towards it. The mariners, however, did not understand the nature of gravitation or the nature of the moon, and, although they might have had a thousand myths and superstitions to draw on, they could never have given a correct (or even close to correct) scientific explanation of the tides. Their inability to explain the phenomenon, however, did not affect their ability to predict it. Many ancient mariners[74] and scientists kept careful track of the moon's position in relation to the tides, and made very rational and accurate predictions of the tides based on their data. This type of prediction, which is based on a strong induction and is empirically verifiable (or falsifiable), should surely not be grouped with the predictions of the psychic or the astrologer[75]. It seems like a scientific prediction, and indeed, I think it is.

Jeffrey's barometer example is another important one in the literature on both explanation and epistemology. In it, we see a falling barometer, and, because we know that this often (or, for the sake of argument we could say always/almost always) precedes precipitation, we know that it will rain soon. We correctly predict the rain, and if anyone were to ask for our evidence, we would simply point to the barometer reading. The difference between this case and the moon example is that in this case there is nothing about our evidence that could be said to *cause* the predicted event. The drop in atmospheric pressure causes both the barometer and the rain to fall, but the barometer

[74] See Samuel Taylor Coleridge for more on ancient mariners; particularly their rimes.
[75] The 'moon and the tides' counterexample is taken from *Four Decades of Scientific Explanation*.,. 47.

has no effect on the rain. It follows that there is no information about the barometer that, in itself, could *explain* the rain. However, I think it is still quite reasonable to say that we have performed a reasonable and scientific prediction in this case, despite the lack of explanatory power in our evidence.

It might be argued that in the barometer case we have background assumptions about air pressure that influence our prediction. This could be the case, but we could certainly imagine having the same predictive power even if our knowledge of air pressure was confused or non-existent. Imagine the person who believes that the mercury in the barometer falls because of an increase in the force of gravity. The gravity pulls on the mercury, and when it gets strong enough it eventually "pulls" the rain out of the clouds. I would want to say that this person might still be in a position to perform an accurate inductive inference, and thus be in a position to make a proper scientific prediction.

Explanations as Essentially Predictive:

Of more crucial importance to my thesis is the question of whether or not explanations are essentially predictive. After all, if explanations that do not explicitly confer upon their explanans a probability of at least 50% are allowed (such as, of course, aleatory explanations), then there will be some explanations that could never have provided us with any predictive power. Indeed, we might have good reason to predict that the negation of the explanans would obtain, but we could still explain it once it happened. Aleatory explanations do not, therefore, satisfy the part of the symmetry thesis that requires an explanation to be predictive. Let us look at some counterexamples to the idea the explanations are essentially predictive.

One form of tertiary syphilis is paresis. Paresis only occurs in those who have been through the primary, secondary and latent stages of syphilis and have not been treated by antibiotics. Whenever a person contracts paresis, the explanation is that they have had untreated (or improperly treated) latent syphilis. Ask any doctor or medical researcher why a given person has contracted paresis and they will give an explanation much like the one above. However, only about 25% of people who have untreated latent syphilis contract paresis so one could never predict that a particular individual would contract the disease. Here we have a non-predictive medical explanation that seems perfectly reasonable from an intuitive notion of scientific explanations.

Evolutionary biology has a great deal of explanatory power. Its primary appeal as a scientific theory is its ability to explain not only the vast array of different organisms on the planet, but also to explain why particular creatures have become dominant within their environment. Certain types of finch become dominant, says the biologist, because of genetic traits that enhance their ability to survive and reproduce. Despite all this, the evolutionary biologist is unable to make even remotely accurate predictions as to what things will be like in the future. Given a certain isolated island at time t that has on it 20 species of thistle, the biologist could not tell us which, if any, of these thistles would become dominant over the others by t+50 years. Indeed, the antecedent probability that *Homo Sapiens* would occupy the niche it now does was probably very small indeed, but here we are, and we have the evolutionary biologist to give us an explanation of our position.

So here we have two counterexamples to the idea that explanations are essentially predictive. Both of them rest on sciences that we would not want to describe as pseudo-sciences, both of which provide us with real and useful statistical explanations of particular facts. However, it might be argued that these explanations are not fundamentally non-predictive, but that they are merely incomplete. In other words, evolutionary biology would be capable of predicting the evolution and adaptation of certain species in an ideal epistemic circumstance, such that we knew everything there is to know about genes, traits, mating patters, weather patters, geographical information, and the like. Our knowledge in all these matters is limited, but it is not limited *in principle*. It is reasonable to think that someday we will have all the relevant facts, and when that day arrives, biology will have the sort of predictive power that the symmetry thesis requires. According to this view, the current evolutionary explanation of life is really just a stop-gap explanation (not really an explanation at all) that is filling in until we can really understand what is going on.

This view of science (that is cannot offer good explanations unless it has ideal epistemic circumstances), is to say the least, troublesome. *All* branches of science suffer from an epistemic shortcoming, so it therefore follows that we have no good (true) explanations *at all*. This is certainly not the case. We cannot completely discount the explanatory theories that science gives us just because human knowledge is destined to be imperfect. In my chapter on aleatory explanations, I give Humphreys' argument against this very point. Furthermore, this view of the world (that with enough information an explanation could be predictive) seems to presuppose determinism, or at least presuppose that only deterministic phenomena are explicable. As I have said many times, this is not an appropriate presupposition for any theory of explanation to make.

To strengthen the point, let me give an example of a current scientific theory that is, according to many (probably most) physicists, non-deterministic[76]. Uranium 238 has a half-life of roughly $6.5 * 10^9$ years. This means that the probability (or *chance*) that a given atom of Uranium 238 will decay during this time period is 0.5. So, the probability that this same atom will decay (thus undergoing alpha-particle emission and becoming an atom of lead) during any 24 hour time period is very, very small indeed. Furthermore, the probability involved here is, according to the quantum physicist, irreducibly statistical. There is, in principle, no way to get any better knowledge of whether a given particle will decay in a given time period than the knowledge we can get by simply watching enough atoms and observing their rate of decay. In other words, the $6.5 * 10^9$ year half-life is not just a "to the best of our knowledge" statistic. It exhausts all of our predictive powers. We simply cannot say anything more precise than "for any $6.5 * 10^9$ year time period, roughly one half of the molecules of a given sample of Uranium 238 will decay." Now say that a given atom a does in fact decay during the 24 hour period of observation. The physicist would explain this phenomenon in terms of the nature of the nucleus of the isotope, and the nature of barrier penetration and alpha-

[76] As I have said before, it is not the purpose of this thesis to argue for either determinism or indeterminism. I will be content to show that there are situations in which determinism is in question. Perhaps quantum events are governed by "hidden variables," but this is not clearly the case, so it would be a philosophically irresponsible to assume determinism without a good argument.

decay. Further, he would add that it was very unlikely that a would decay during the day in question, but would nevertheless say that he understands the process through which the decay occurred, and might add that the improbable can be expected to happen sometimes (although at no *particular* time). In fact, if the improbable didn't happen *occasionally*, it wouldn't be *improbable,* but rather *impossible*[77].

So here is a case of an explanation that is *essentially* statistical, and that cannot be reduced to any more basic deterministic laws. It seems to be a good explanation in that the explanans give us a good understanding of how and why the atom decayed, but it certainly has no predictive power.

It would appear that there are many cases in diverse scientific fields in which we have explanations that are inherently non-predictive. In some of these (like the syphilis-paresis or evolutionary biology cases), it might seem that our statistical explanation is merely taking the place of some "better" deterministic explanation, while in others (like the Uranium 238 case), the probabilities involved are irreducible. Together with the examples of non-explanatory prediction discussed earlier, it is clear that we have inadequate intuitive support for the symmetry thesis. Such a thesis would, if true, show many of our more convincing scientific theories and explanations to be in error, despite the fact that they are accepted by so many scientists, and so many philosophers of science (like Hempel himself) and, to quote David Lewis, "an analysis that imputes widespread error is prima facie implausible. Moreover, it is dishonest to accept it, if you

[77] That is, physically impossible, not logically impossible. But perhaps even this usage is misleading. At any rate, it might not be called *impossible* in any sense, but neither would it be called improbable in the normal sense.

yourself persist in the "error" when you leave the philosophy room."[78] We do in fact accept non-predictive explanations and non-explanatory predictions from science every day. We would need a much better argument than the ones Hempel gives to believe that we are wrong to do so.

The Desire for Determinism

As I have mentioned, it seems to me that both van Fraassen's view that explanations should be contrastive and Hempel's view that they should be predictive stem from a tacit assumption that the universe is, when all is said and done, deterministic. I doubt that either of these philosophers would explicitly hold this view, but, as Peter Railton puts it, "Some people who do not hold the doctrine of determinism are nevertheless held by it."

It seems commonplace for most of us to forget, when we leave the philosophy room or the laboratory, that determinism is not the obviously correct state of things. The world that presents itself to our senses on an everyday basis *seems* quite regular and deterministic, and sometimes this commonplace notion makes us forget the last one hundred years of physics. The advantages of Humphreys' model of explanation should not be overlooked simply because they fail to meet with our anachronistic desire that things be rigid and deterministic. If our naïve intuitions are correct, and the world is deterministic, then aleatory explanations will help us understand it. If some of the prevailing scientific views of indeterminism are correct, then, presumably, aleatory

[78] David Lewis. "Causation." p. 175

explanations will help us to understand those. The point is that the question is left to the scientists to decide[79].

[79] I know that I have repeated this criterion for the adequacy of an explanatory account time and time again, and it might seem somewhat repetitive (it is), but when I look at the literature on scientific explanation, this is *the* most important difference between the various accounts. Many of the debates and disagreements regarding scientific explanations either explicitly or implicitly turn on this issue, so I want to be as clear as possible. If that means a bit of repetition whenever this point recurs, then so be it.

CHAPTER VII

ALEATORY* EXPLANATIONS

I have spent the bulk of this thesis arguing that Paul Humphreys' theory of aleatory explanations is the correct one on which to model our scientific explanations. The reasons for this are many and diverse, and have all been given in the preceding chapters. In this chapter I want to propose a way in which the theory can be improved over the presentation that Humphreys gives of it. This is not to say that aleatory explanations have any sort of fatal shortcoming, only that improvement is possible.

The Problem

As mentioned above, Paul Humphreys offers the following sufficient conditions on one explanation being *better* than another. For any two true explanatory accounts "Y because Φ, despite Ψ" and "Y because Φ', despite Ψ,"

> If Φ _ Φ' and Ψ = Ψ', then the explanation of Y by Φ' is superior to that given by Φ. If Φ = Φ' and Ψ _ Ψ' then again Φ' gives a superior explanation, in the sense that it is more complete.[80]

I have already said that he would almost certainly, based on the style of these conditions, accept that when Φ _ Φ' and Ψ _ Ψ', then the explanation given of Y by Φ' is superior to that given by Φ. So let us take these three conditions as sufficient conditions on one explanation of Y being superior (better) to another explanation of Y.

---

[80] "Scientific Explanation: The Causes, Some of the Causes, and Nothing but the Causes" p. 296.

Obviously these are not necessary conditions on the comparative quality of an explanation, and so much the worse for any theory that says they are. Surely "because of the speed of the car and the sand on the road, despite the functioning brakes" is a better explanation of a car's skidding off the road than "because of the gravitational force exerted on that car by Alpha Centauri," even when both are true. But this example does not meet any of the above conditions. So we know that the conditions cannot be necessary.

Furthermore, the conditions cited do not tell us when an explanation is "good" or "bad," only when it is "better" or worse." That both explanations are true is assumed, as a *false* "explanation" is neither better nor worse than any other explanation—it is not an explanation at all, but beyond that, all we can determine from the form alone is the *relative* quality of the explanation. Science may be able to come up with objectively good and bad explanations, but we will actually have to know something about science to make this determination, not just about the form of the explanation.

So these conditions as given are clearly not all that strong, and, it would seem, they match our intuitions about the quality of a causal explanation quite well. After all, when causes are in question, if explanatory account A can cite all the causes that account B cites, plus some, then A must be a better explanation that B, at least in the sense that it is more complete. That is, it gives you a better understanding about why the explanandum should or should not have happened.

However, I believe I have a counterexample. Imagine that I set up a domino-topple consisting of 100 separate dominos. Further, imagine that the domino-topple is completely isolated, completely deterministic, and arranged such that every domino hits one and only one other domino in a straight line. Further, imagine that they are spaced in such a way that if any of my dominos were removed (other than the last one), then the chain would be broken and the last one would not fall. Okay, now imagine that I tip the first domino and the line topples over in the standard way, with the result that the last domino falls. Now take P to be the event described by "The last (100th) domino fell," and Q to be the event described by "The first domino was tipped over."[81] The following aleatory explanation seems appropriate:

(A) P because Q.

It is in the canonical form and it correctly identifies a contributing cause. It is therefore a true explanation. But now imagine that R is the event described by the sentence "The second domino fell." Consider the following explanation:

(B) P because R.

Now, it would seem that we have yet another perfectly good explanation. Certainly R was a cause of P by any of the standard notions of causation (either counterfactually "had R not occurred, P would not have either" or on a frequency theory of causation "whenever R falls, P does also"). In fact, it would appear that Q and R both confer the exact same likelihood upon the occurrence of P (in this case, that likelihood is one). If either explanation is superior, I would be inclined to say it is A, simply because the

[81] It is important the P and Q are not the sentences themselves, but rather the events described by the

cause Q, being temporally prior to R, and being the direct, deterministic cause of R, gives us more information about the events that led to P. To use Humphreys' terminology[82], the explanation A is a "superior explanation, in the sense that it is more complete." Of course there is no way of determining their relative quality based on what Humphreys has given us with regard to purely *formal* superiority (neither Φ is a proper subset of the other), so we are left to our science or our intuitions, but this is no problem. The problem arises when we consider the explanation:

(C) P because Q and R.

C is, according to Humphreys' account, a *better* explanation than either A or B, since it cites more of the causes involved it is superior on the aleatory account. But I do not think this is obvious at all. In fact, I think that A, and C are all equally good explanations of the event P, while both are superior to B, since they give more of the causal history of P.

My reason for saying this is that despite the fact that C cites more causes than either A or B, it does not give us any better understanding of the event P than they do. Since R is causally implied by Q (we are assuming), that is, since all of the effects of R must take place whenever Q occurs (or, counterfactually, Q □→ R), saying that R occurred really adds nothing to the account over and above saying that Q occurred.

---

sentences. After all, Φ and Ψ are sets of causes, not sentences.

[82] Though I do not know for sure whether or not Humphreys' intuitions would agree with mine in this particular case

Though more causes are cited in C, the second cause is, in a sense, uninteresting and uninformative.

Just to make the case a bit more convincing, let me modify it a little. Imagine S is the state of affairs prior to the toppling of the first domino. A description of S would look something like "There are 100 dominos lined up such that if the first one falls, each successive one will fall deterministically, etc., resulting in the last (100^{th}) domino falling." The following three explanations are both true:

(A') P because Q and S.
(B') P because R and S.
(C') P because Q and R and S.

Here I think it is obvious that A' is better than B', again, in the sense that it is more complete. By stating that Q occurred, we are saying something over and above the fact that R occurred. After all, we could easily imagine R having occurred in Q's absence, say, if a gust of wind had tipped over the second domino but not the first. But it seems equally clear to me that C' does not say anything over and above what is said by A'. Each explanation is equally good. Just to drive the point home, lets look at the English translations of A' and C'. A' says:

The 100^{th} domino fell over because it was the 100^{th} out of 100 dominos, arranged such that if the first one were to fall all the others would be deterministically sure to fall in succession, right down to the 100^{th} one, and the first one did in fact fall.

And C' says:

The 100th domino fell over because it was the 100th out of 100 dominos, arranged such that if the first one were to fall all the others would be deterministically sure to fall in succession, right down to the 100th one, and the first one did in fact fall, followed by falling of the second domino.

The final clause of the second explanation adds absolutely nothing to the explanation that is not already present, and hence, I do not think we would be justified in calling C' a better explanation than A'. If there are those who disagree, and see C' as somehow a more complete account, giving a better understanding of the explanandum, I do not know if I can do more to convince them. Regardless, my intuitions strongly favor the claim that A' and C' are equally good explanations of the event P.

The Solution

So we clearly have a problem with the account of aleatory explanations as it stands. Though Humphreys model is the best out of all those I have examined, it needs some modification regarding its handling of the comparative quality of explanations. I think I can give just such a modification.

I said that the problem with regarding C' to be a better explanation than A' is that the account given by C' adds nothing to the account given by A'. The reason that this is the case, it seems, is that the sentences describing Q and S logically entail the sentence describing R. This should be clear if you go back and read the spelled-out English version of the explanation C'. From the proposition, "There were 100 dominos, arranged such that if the first one were to fall all the others would be deterministically sure to fall in succession, right down to the 100th one," and the proposition "The first domino fell," it should be a simple matter of logic to infer that the second domino fell as well. If we understand 'determinism' and the notion of 'falling in succession', then the inference should be pretty straightforward:

For all D_n, where 1[n[100, if D_{n-1} falls, then D_n falls.
D_1 falls.
Therefore, D_2 falls.

So it is clear that R (which is just the conclusion of this argument) is logically entailed by S and Q (which are the premises). One might object that such an inference assumes determinism, and hence it is not true that the falling of domino one entails the falling of domino two. But this, while true, does nothing to block the inference, since the additional premise S is involved, in which determinism is explicitly assumed. Sure, there are possible worlds in which the first domino falls and the second does not, but these are not worlds in which the premises of the argument (which are just statements of the causes in our explanation) are true, so they need not concern us. Furthermore, my earlier qualms with essentially deterministic explanations still obtain, but that does not prevent me from stipulating that determinism is true in a particular instance.

So, it would appear that the real problem with thinking that explanations A' and C' could be better or worse than one another is that an accurate description of the causes they mention shows that the set Φ in each explanation is logically entailed by the set Φ in the other. This leads me directly into how to solve the problem with Humphreys' account. The sufficient conditions on the superiority of an explanation Φ' should be amended as follows:

If Φ _ Φ' and Ψ = Ψ', then the explanation of Y by Φ' is superior to that given by Φ, unless a complete description of the elements of Φ entails a description of the elements of Φ'.
If Φ = Φ' and Ψ _ Ψ' then again Φ' gives a superior explanation, in the sense that it is more complete, unless a complete description of the elements of Ψ entails a description of the elements of Ψ'.
If Φ _ Φ' and Ψ _ Ψ', then the explanation of Y by Φ' is superior to that given by Φ, unless a complete description of the elements of Φ and Ψ entails a description of the elements of Φ' and Ψ'.

Although I have only discussed the problem of logical entailment as it regards contributing cases, it should be obvious that a similar problem confronts counteracting causes, and that it can be solved in the manner described above. Further, you will notice that I have made explicit the condition (which I am sure Humphreys would assent to) that when Φ and Ψ are *both* proper subsets of Φ' and Ψ', respectively, then the explanation given by Φ' is superior to that given by Φ (plus, of course, the unless-clause that I have added).

I contend that the above modification of Humphreys' model of aleatory explanations is necessary in order for it to properly preserve our intuitions regarding the relative quality of scientific explanations. I cannot say, however, whether or not Humphreys would agree with my intuitions or my conclusions in this matter, so from here on, I will refer to my modified version of his model as aleatory* explanation, and let 'aleatory explanations' continue to refer to the view Humphreys himself presents us with.

CHAPTER VIII

CONCLUSION

In conclusion, I feel that I have given a number of convincing reasons why the traditional models of scientific explanation are inferior to the account of aleatory explanations given by Paul Humphreys. Aleatory explanations are causally oriented, non-inferential linguistic devices that rigidly determine what counts as true and false explanations in science. They allow for the essential incompleteness of scientific discourse, and hence no explanation that fits the model will ever be *complete* and yet many of them are still *true,* and good sense can be made of one such explanation's being superior to another. Further, they have no requirement of maximal specificity or other such restriction that would block all by the most trivial explanations. Aleatory explanations are not essentially contrastive, nor are they predictive, but this is a distinct advantage, since they do not thereby imply determinism.

However, I have pointed out that there is a slight difficulty involved regarding the sufficient conditions that Humphreys gives of one explanation's being superior to another. I corrected this problem and titled the revised account aleatory* explanation. Since it would appear that aleatory* explanations have all the advantages of aleatory explanations, but one less disadvantage, I contend that the revised account is superior to all of the standard models of scientific explanation. It is, I believe, at least a very close approximation of the *correct* view of scientific explanation, corresponding both to our intuitions about explanations and to the actual practice of science.

REFERENCES

Achinstein, Peter. “The Symmetry Thesis.” In *Science, Explanation, and Rationality: Aspects of Carl G. Hempel,* edited by James Fetzer, 167-85. Oxford: Oxford University Press, 2000.

Bromberger, Sylvain. “An Approach to Explanation.” In *Analytical Philosophy: Second Series*, edited by R.S. Butler, 72-105. Oxford: Blackwell Publishers, 1962.

Coffa, Alberto. “Hempel’s Ambiguity.” *Synthese* 28 (1974): 141-163.

Fetzer, James, ed. *Science, Explanation, and Rationality: Aspects of Carl G. Hempel.* Oxford: Oxford University Press, 2000.

Van Fraassen, Bas. *The Scientific Image.* Oxford: Oxford University Press, 1980.

Hempel, Carl. “Aspects of Scientific Explanation” In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: The Free Press, 1965.

—“Deductive-Nomological Versus Statistical Explanation.” In *The Philosophy of Carl G. Hempel: Studies in Science, Explanation and Rationality,* edited by James Fetzer, 87-145. Oxford: Oxford University Press, 2001.

—“Maximal Specificity and Lawlikeness in Probabilistic Explanation.” In *The Philosophy of Carl G. Hempel: Studies in Science, Explanation and Rationality,* edited by James Fetzer, 146-64. Oxford: Oxford University Press, 2001.

Hempel, Carl, and Robert Oppenheim. "Studies in the Logic of Explanation." *Philosophy of Science* 15, No. 1 (1948): 135-75.

Humphreys, Paul. "Aleatory Explanations." *Synthese.* 48 (1981): 225-32.

—“Scientific Explanation: The Causes, Some of the Causes, and Nothing but the Causes.” In *Minnesota Studies in the Philosophy of Science,* edited by Philip Kitcher and Wesley Salmon, 271-295. Minneapolis: University of Minnesota Press, 1989.

—"Why Propensities Cannot be Probabilities." *The Philosophical Review,* 94, No. 4 (1985): 557-570.

Jeffrey, Richard. “Statistical Explanation vs. Statistical Inference.” In *Essays in Honor of Carl G. Hempel,* edited by Nicholas Rescher, 104-113. Dordrecht: D. Reidel Publishing, 1969.

Kitcher. Philip, and Salmon, Wesley, eds. *Minnesota Studies in the Philosophy of Science*. Minneapolis: University of Minnesota Press, 1989.

Lewis, David. “Causal Explanation.” In *Philosophical Papers, Vol. 2,* 214-240. Oxford: Oxford University Press, 1986.

—“Causation.” In *Philosophical Papers, Vol. 2,* 170-214. Oxford: Oxford University Press, 1986.

Railton, Peter. “A Deductive Nomological Model of Probabilistic Explanation.” In *Theories of Scientific Explanation*, edited by Joseph Pitt, 119-135. Oxford: Oxford University Press, 1988.

Rescher, Nicholas, ed. *Essays in Honor of Carl G. Hempel*. Dordrecht: D. Reidel Publishing, 1969.

Salmon, Wesley. *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota press, 1989.

—“Statistical Expanation and Causality.” In *Theories of Scientific Explanation*, edited by Joseph Pitt, 75-118. Oxford: Oxford University Press, 1988.

VITA

Peter A. Sutton
100 Graham Rd. #13A
Ithaca, NY 14850
(607) 257-7936
pas45@cornell.edu

Education

1996-1999 Texas State University
Bachelor of Arts, History & Philosophy
San Marcos, TX 78666

2000-2002 Texas A&M University
Master of Arts, Philosophy
College Station, TX 77843

2002-Present Cornell University
Ph.D. program in Philosophy
Ithaca, NY 14950

Teaching Experience

2000 – 2002 Teaching Assistant: Introduction to Logic and Professional Ethics
Department of Philosophy, Texas A&M University
314 Bolton Hall
College Station, TX 77843-4237
(979) 845-5660 (Robin Smith, Department Head)

Summer 2002 Instructor: Introduction to Logic
Department of Philosophy, Texas A&M University

2002-Present Teaching Assistant: Introduction to Logic, Logic and Computation, Introduction to Philosophy, and Philosophy of Sport
The Sage School of Philosophy, Cornell University
218 Goldwin Smith Hall
Ithaca, NY 14850
(607) 255-3687