A Metaphysically Neutral Theory of Singular Scientific Explanation

Michael E. Cuffaro

A Thesis

in

The Department

of

Philosophy

Presented in Partial Fulfilment of the Requirements

for the Degree of Master of Arts (Philosophy) at

Concordia University

Montréal, Québec, Canada

August 2008

Your file Votre référence
ISBN: 978-0-494-45283-7
Our file Notre référence
ISBN: 978-0-494-45283-7

ABSTRACT

A Metaphysically Neutral Theory of Singular Scientific Explanation

Michael E. Cuffaro

Modern philosophical debate on explanation began with Hempel and Oppenheim's 1948 article: *Studies in the Logic of Explanation*. Hempel and Oppenheim's view: their Deductive-Nomological model of explanation for deterministic phenomena, and the Inductive-Statistical model of explanation for indeterministic phenomena, were for many years the received view on the subject. That view has since fallen out of favour, however, and in its wake, many alternative models have been proposed. Of these, Peter Railton's Deductive-Nomological-Probabilistic model represents an exceptionally promising proposal. In this thesis, I argue that while we should accept the essentials of Railton's account, we should not endorse it in its totality. Railton's account of explanation forces us to ascribe to a hard form of reductionism. It also forces us to accept a propensity interpretation of probability and a causal interpretation of irreducibly indeterministic phenomena. In this thesis I argue that one should not have to engage in these metaphysical debates - that an account of explanation should be neutral with respect to these issues.

ACKNOWLEDGEMENTS

I offer my sincere thanks to my advisor, Dr. Gregory Lavers, for his helpful and incisive criticisms of my earlier drafts of this thesis, for his encouragement and support, for his sober advice, and for proving that I am not the only one who thinks that logic can be fun. I also thank Dr. Murray Clarke for reinforcing in me the belief that philosophy should make a difference and that philosophy should be relevant, Dr. Pablo Gilabert for challenging me and for demanding from me nothing but my best, Dr. Dennis O'Connor for helping me to find connections where I would never have thought to look for them, and Dr. Vladimir Zeman, who more than anyone has convinced me that philosophy is worth doing.

I offer my thanks, also, to the rest of the faculty and to the staff of the Philosophy department for making my experience at Concordia a pleasant and productive one, to my fellow students, both graduate and undergraduate, from whom I have probably learned the most in this past year; lastly I thank my friends and I thank my family, with whom I have always been and always will be philosophising.

Table of Contents

Introduction

Giving a precise account of the nature and purpose of scientific explanation is a challenge that is about as old as philosophy itself, but modern philosophical debate on explanation began in the 20^{th} century with Hempel and Oppenheim's Deductive-Nomological model of explanation for deterministic phenomena. That model, and Hempel's later Inductive-Statistical model for indeterministic phenomena, were for many years the received view on the subject. On that view, explanations are arguments; the purpose of an explanation is to convince one that the fact which it explains follows either deductively or inductively from the bulk of our knowledge.

Hempel and Oppenheim's view has since fallen out of favour, however, and in its wake, many alternative models have been proposed. Of these, I believe that Peter Railton's Deductive-Nomological-Probabilistic model represents an exceptionally promising proposal. Railton's conception of explanation is an ontic one; i.e., he holds that a good explanation is one that demonstrates the mechanisms and processes which, according to our best theories, truly give rise to a phenomenon. One of the many virtues of Railton's account is that while it places a strong emphasis on the role of causes in explanations, it allows that some explanations may not be causal in nature. I agree, and I will argue that any account of explanation must be able to deal with non-causal cases - and that Railton's model satisfies this criterion. But while I accept the essentials of Railton's account, I do not endorse it in its totality. Railton's account of explanation forces us to ascribe to a hard form of reductionism. It also places too much emphasis on the appeal to causes, forcing us to accept a propensity interpretation of probability and a causal interpretation of irreducibly indeterministic phenomena. I will argue that one should not have to engage in these metaphysical debates when giving an account of scientific explanation. Ideally, I will argue, an account of explanation should be neutral with respect to these issues.

I do not believe that these metaphysical issues are either meaningless or unimportant, however. I do not believe that metaphysics is nonsense and I do not desire to see it banished from philosophy (or even from science). Yet there are times when we must put aside these questions. I believe it is possible to give an *account* of explanation that assumes the answers to none of them and that is able to make room for the answers to them all. Railton, although himself a realist, has put forth a model of explanation that is capable, in my eyes, of doing just this. Yet Railton's realism has intruded too much. Thus I will argue that we should accept Railton's model, minus his metaphysics.

A final note before I begin: as the title of this thesis implies, my focus here will be on *singular* explanation; that is, explanation with regard to singular facts: the collapse of a bridge, a car accident, the gravitational pull of the sun, and so on. I will not extend my discussion to models, such as those put forward by Friedman, Kitcher, and others, for explanations of general regularities (where one or more specific laws are subsumed under more general ones).[1] I do this merely for lack of space. The subject of scientific explanation is a broad one. There is already very much to discuss with regard to singular explanation; moreover, I believe that the model of explanation that I present is capable of being extended to account for explanations of both singular facts and general regularities. But that is not something that I wish to defend in this thesis.

[1] See, for example, Friedman 1974, Kitcher 1989, for these unificationist accounts of explanation.

Chapter 1

Explanations of Deterministic Phenomena

1.1 The Deductive-Nomological Model

Law Statements and Theory Statements

An explanation of a singular fact (e.g., an eclipse of the sun, the salt content of a body of water) can be characterized as follows: it consists of a set of sentences subdivided into (a) an *explanandum* statement: a sentence describing the fact to be explained, and (b) an *explanans*: a set of sentences that describe the relevant factors which account for that fact. With regard to the explanans, on Hempel and Oppenheim's view it must always include at least one 'general law': either a fundamental law, a derivative law, or a theory. We can explicate the concept of a *fundamental* law in a language L as follows:[1]

> (7.3a) S is a fundamental lawlike sentence in L if S is purely universal; S is a fundamental law in L if S is purely universal and true.

[1] Hempel 1965, 272. The numbering here (7.3a, 7.3b, etc.) corresponds to the numbering of definitions given in Hempel and Oppenheim's original article.

Fundamental lawlike sentences are universally quantified sentences, e.g., of the form: $(\forall_x)(B_x \supset C_x)$, for which we impose no restriction on the domain. Fundamental *laws*, on the other hand, are just those fundamental lawlike sentences that are also true. Next, we define the concept of a *derivative* law in L as follows:

> (7.3b) S is a derivative law in L if (1) S is essentially, but not purely, universal and (2) there exists a set of fundamental laws in L which has S as a consequence.

Derivative laws are universally quantified sentences whose domain is restricted (this is what is meant by saying that they are not 'pure'). Newton's laws of motion are examples of derivative laws, for they are not valid for very small scales, for very high speeds, or for very strong gravitational fields. They are derivable from the fundamental laws, however, so they do not lose their law status in L. Thus,

> (7.3c) S is a law in L if it is a fundamental or a derivative law in L.

Before continuing, let me point out that (7.3a) does not rule out tautologies (i.e., laws of inference). While Hempel and Oppenheim require that every explanation has empirical import, they do not interpret this requirement so strongly so as to impose it on every statement in the explanans. They require only that the explanans as a whole (in other words, at least one of its law statements) has empirical import.

We come, finally, to what Hempel and Oppenheim call *theories*:[2]

> (7.4a) S is a fundamental theory if S is purely generalized and true.
>
> (7.4b) S is a derivative theory in L if (1) S is essentially, but not purely, generalized and (2) there exists a set of fundamental theories in L which has S as a consequence.
>
> (7.4c) S is a theory in L if it is a fundamental or a derivative theory in L.

[2] *Ibid.*, 272.

The difference between law sentences and theories is that the latter can include existential as well as universal quantifiers (sentences such as: $(\forall x)(\exists y)(P_{xy})$; e.g., 'every comet has a tail').[3].

D-N Arguments

Explanations, for Hempel and Oppenheim, are arguments. They involve the subsumption of a fact under a law or a set of laws. We call an explanation *deductive-nomological* or 'D-N' when all of the laws involved are deterministic and satisfy requirements 7.3 - 7.4.[4] Later, I will consider the case of so-called *inductive-statistical*, or 'I-S' explanation involving one or more 'statistical' laws,[5] but for now I will limit the discussion to the deterministic case. We can represent a D-N explanation, according to Hempel and Oppenheim, by way of the following schema:

$$\begin{array}{ll} explanans & \left\{ \begin{array}{l} C_1, C_2, ..., C_k \\ \underline{L_1, L_2, ..., L_k} \end{array} \right. \\ explanandum & \qquad E \end{array}$$

Here, $C_1, C_2, ..., C_k$ represent the initial conditions, while $L_1, L_2, ..., L_k$ are our general law statements. Together, they form the explanans (the premises). E, the explanandum statement (the conclusion of the argument), follows deductively from the explanans.[6] It is important to note that the general law statements, $L_1, L_2, ..., L_k$, cannot be omitted from an explanation; we do not permit the explanandum statement to follow directly from the initial conditions alone.[7] More formally, we can put this as follows:[8]

[3]Hempel and Oppenheim do not actually specify if there is a particular order in which the quantifiers need to occur. Also, it is an open question whether theories require universal quantifiers at all. See Salmon 1989, 19, for a discussion of this

[4]Cf. Above.

[5]A statistical law can be seen as a less stringent version of a law such as: $(\forall x)(P_x \supset R_x)$ (See Hempel 1965, 377). Most statistical laws are reducible, at least in principle, to deterministic laws; however, some (the laws of quantum mechanics) are widely held to be irreducibly indeterministic.

[6]Cf. Hempel 1965, 336.

[7]Cf. *Ibid.*, 247-248.

[8]*Ibid.*, 273.

(7.5) An ordered couple of sentences, (T, C), constitutes a potential explanans for a singular sentence E only if

(1) T is essentially generalized and C is singular

(2) E is derivable in L from T and C jointly, but not from C alone.

Finally, we say that a *potential* explanation is *true* iff T is a theory (i.e., iff T is also true). For example, consider an astronaut on a lunar walk. At time t_0, she lets go of a moon rock from a height of 3 metres. Now we would like to explain the fact that the rock hits the ground at $t_0 + 3s$, and we can do it as follows:[9]

$C_1 : v_{t_0} = 0m/s$	(the rock is initially at rest)
$C_2 : \Delta x = 3m$	(the distance to the ground is $3m$)
$C_3 : M = 7.3477X10^{22}Kg$	(the mass of the moon)
$L_1 : F = G\frac{M}{r^2}$	(the law of universal gravitation)
$L_2 : \Delta x = v_0 t + \frac{1}{2}at^2$	(law for motion in one dimension)
E : The rock hits the ground at time t_0 + 3s.	

The Thesis of Structural Identity

Part of the motivation for representing explanations as arguments is that, for Hempel,[10] explanations should double as potential predictions. Explanations should not only increase our understanding of observed phenomena; they should allow us to use that understanding to extend our knowledge to (currently) unobserved phenomena as well. This is what Hempel calls the *thesis of structural identity*. Since explanations are arguments, the explanandum always follows from the explanans. Therefore if we know that the conditions and laws explaining the occurrence of some phenomenon obtain, we can use those same conditions and laws to predict future occurrences of the phenomenon. For example, the explanation of a particular eclipse of the sun can be used to predict future eclipses. Ex-

[9]Below, I use the universal law of gravitation in order to derive the rock's acceleration, $1.6m/s^2$, as it falls towards the ground (I have omitted the actual derivation). Note that we can ignore the mass of the rock in the calculation as it is very small compared with the moon.

[10]From now on I will refer exclusively to Hempel; while Oppenheim co-authored their 1948 article, the later elaborations and defences of both the D-N and I-S models were due to Hempel alone.

planations and predictions are thus symmetrical: an explanation is a potential prediction; a prediction (in D-N or I-S form[11]) is a potential explanation.

But now consider the following case.[12] Imagine a flagpole standing in a field on a level stretch of ground. Given its height, along with the current position of the sun and the relevant laws of physics, we can use a D-N argument to infer that (and hence explain why) it casts a shadow of a particular length. So far, so good; this is a perfectly acceptable explanation of the shadow's length, and it is clear that the D-N argument in this case can also be used to predict the length of other shadows given similar circumstances. But now note that we can also construct a D-N argument inferring the height of the *flagpole* from the length of its shadow.[13] But while this argument serves perfectly well for *predicting* the height of the pole (say, if it was obscured somehow), one would likely get strange looks if one said that one could *explain* the height of the flagpole from the length of its shadow. We can easily imagine similar examples (e.g., a barometer is good for predicting the weather, but one would be hard-pressed to argue that the barometer actually *explains* the weather).

What these cases show is that often there is an asymmetrical cause-effect relationship between the facts cited in the explanans and the explanandum event, and that in these cases it is impossible to transform a D-N prediction into a D-N explanation. I think this objection is dead on. However, we should note that it is possible to give Hempel's thesis both a strong and a weak reading. Hempel's thesis, in fact, consists of two sub-theses: "... *the thesis of structural identity amounts to* the conjunction of *two sub-theses*, namely (i) that *every adequate explanation is potentially a prediction* ... (ii) that conversely *every adequate prediction is potentially an explanation*".[14]

What the flagpole example shows is that sub-thesis (ii) is not universally true. Hempel is willing to admit this, however, for he writes: "I will argue that the first sub-thesis is

[11]There is no requirement on predictions that they be given in I-S or D-N form. Hempel allows that some predictions can be given on the basis of observation alone without appeal to laws. Cf. *Ibid.*, 368.

[12]Cf. Bromberger 1969.

[13]That is, by switching one of the initial conditions for the explanandum.

[14]Hempel 1965, 367.

sound, whereas the second one is indeed open to question".[15] Now even if sub-thesis (ii) is false, I do not think that Hempel loses much. Clearly, sub-thesis (i) is the more important of the two: the D-N model is primarily a model of explanation, not of prediction. The thesis of structural identity is intended to highlight a certain fact about ideal explanations: that they are fruitful, in the sense that they serve to *expand* our knowledge of unobserved phenomena. This is unaffected by the falsity of sub-thesis (ii). As for sub-thesis (i), it seems impossible to refute it, for the explanandum-statement is, after all, logically entailed by the explanans. Salmon (a vocal critic of Hempel's) concedes this point.[16]

Some have argued against sub-thesis (i), however. Scriven, for instance, asks us to consider the case of a collapsed bridge.[17] Imagine that upon observing a collapsed bridge, we make the following hypothesis: the bridge collapsed either (a) as a result of metal fatigue, or (b) as a result of excessive load. Imagine further that it would have been impossible to know about the metal fatigue before the actual collapse of the bridge. Now we explain the collapse as follows: we know that excessive load was not a factor (the bridge collapsed in the middle of the night), and thus it follows that since the bridge *did, in fact, collapse*, that the collapse was caused by metal fatigue. But note that apart from the bridge's actual collapse there is no direct evidence that metal fatigue was strong enough to cause it; thus we cannot turn this explanation into a (non-circular) prediction.

But on Hempel's behalf I will say that first, it is stretching things to say that there *could not* have been any evidence of excessive metal fatigue before the actual collapse of the bridge. As Hempel writes: "the impossibility appears to be rather a practical and perhaps temporary one, reflecting present limitations of knowledge or technology".[18] Second, even if before the collapse our knowledge was insufficient to predict it, what has been *learned* (by means of an explanation resulting from an investigation into causes of the collapse) can indeed be used to predict the collapse of some *other* similar bridge. Thus I do not believe

[15] *Ibid.*, 367.
[16] Salmon 1989, 49.
[17] Scriven 1962, 181-187.
[18] Hempel 1965, 371.

that Scriven's counter-example represents a serious challenge to sub-thesis (i). For D-N explanations, at least, I conclude we can grant this sub-thesis to Hempel.

However, there still seems to be something amiss. For to say that some D-N arguments are not explanations does not tell us why this is so. Schematically, both versions of the flagpole example above are identical; yet one counts as an explanation and one does not. A satisfactory account should give us some reason, apart from our intuitions, to distinguish between the two. It should give us sufficient, not just necessary, conditions for what counts as an explanation. It seems obvious that what is missing in the above cases is some sort of causal account. In fact, the problem of asymmetry has led some to assert that what we require in explanation is an explicit appeal to causes. For instance, Salmon writes:

> When we come to the second question, regarding temporal asymmetry, we cannot avoid raising the issue of causation ... The time has come, it seems to me, to put the "cause" back into "because." Consideration of the temporal asymmetry issue forces reconsideration of the role of causation in scientific explanation, and of the grounds for insisting that occurrences are to be explained in terms of antecedent causes rather than subsequent effects.[19]

1.2 Causal Aspects of Explanation

Salmon is one of a few philosophers of science who hold that explanations must primarily be construed as causal accounts: 'Why did E occur?' 'Be*cause* of C'.[20] Salmon's own Causal-Mechanical (C-M) account of explanation is ontic; i.e., he holds that to give an

[19]Salmon 1984, 96-97.

[20]Others are, for instance, Ruben, who actually takes the foregoing sentence as a 'model'; explanations, for Ruben, are simply single sentences, of the form 'c is the cause of e' (Ruben, 732). Humphreys 'aleatory' view is somewhat similar:

> If one wishes to request an explanation, the canonical form will be: "What is the explanation of Y in S at t?" An appropriate explanation will be "Y in S at t [occurred, was present] because of Φ, despite Ψ" where 'Y', 'S', 't' are terms referring to, respectively, a property or change in property, a system, and a time; 'Φ' is a (nonempty) list of terms referring to contributing causes of Y; and 'Ψ' is a (possibly empty) list of terms referring to counteracting causes of Y (Humphreys 1989, 287).

explanation of a phenomenon is to give a literally true story about the mechanism by which a phenomenon arises (which for him is always causal). But if we take an ontic view of explanation, and if causality is to play such a crucial role, then we cannot get away without explicating the concept of causality. We must either define causality in terms of other, more primitive concepts, or else we must show that such a reduction is impossible. We cannot just use this concept naïvely. Salmon, to his credit, does undertake such an analysis. On his highly influential account, causality is explicated in terms of causal *processes*, where a causal process is a process that *transmits* something between two points on its world-line.[21] Examples of causal processes are mass particles, radio waves, photons, even you and I. On the original version of Salmon's account, for a process to qualify as genuinely causal it must be capable of transmitting a *mark*. By a 'mark', Salmon means *information* - a signal. Thus a piece of paper, for instance, qualifies as a causal process since it is capable of transmitting a signal (a letter that I write to you); the radio waves emitted from your local FM station are likewise causal processes. We need not limit the examples to these anthropocentric ones: anything that is capable of transmitting a change in its structure from one point in its world line to another is a causal process.

Salmon's inspiration for the mark transmission theory comes from Einstein's special theory of relativity, according to which nothing can travel faster than the speed of light in a vacuum. What kind of 'thing' we mean in this context needs to be qualified, however, for certain 'things' do, in fact, move faster than light. For instance, consider a spotlamp rotating, once per second, on its base in the middle of a large stadium. Now the *spot* that it casts on the wall of the stadium does one full circuit of the stadium at exactly the same rate as the rate of rotation for the spot*lamp* (once per second) - and this is independent of the stadium's size. If we do the calculation, we will find that when the radius of the stadium is approximately 50,000km, the spot moves around the stadium at a rate equal to the speed of light. Increase the radius of the stadium yet further, and the spot is now actually travelling

[21] In physics, the world line of an object is the unique path that something takes through four-dimensional spacetime throughout its history.

faster than light.[22] But this does not amount to a contradiction of special relativity, for the precise formulation of that principle is that no *signal* can travel faster than light. The spot on the wall, unlike the beam of light emanating from the spotlamp, cannot transmit information. In Salmon's terminology, the spot is a *pseudo-process*. Now both pseudo-processes and genuine causal processes have a certain structure; the difference, however, is that pseudo-processes are incapable of transmitting this structure from one point on their world-line to another.

Consider, also, the following example of a car and its shadow:[23] Salmon asks us to imagine a car travelling down the road, with its shadow travelling along with it. Suppose, now, that the car hits a wall. After the collision, the car will carry the marks of the collision (scratches, dents, etc.) with it. But suppose that only the shadow hits a wall. In this case, the shadow will be momentarily deformed, but it will take its former shape again as soon as it is clear of the wall. Similarly, suppose I place a large quantity of dynamite into the trunk of the car and ignite it; the car will be obliterated; I will have blasted the car out of existence. But if just the shadow is obliterated (as the car travels, say, through a dark tunnel), it will pop back into existence as soon as the car re-emerges. Now if a causal process is defined as one that is capable of transmitting its own structure, then the mark criterion is what enables us to differentiate a causal from a pseudo-process. For a mark represents a *change* in that structure. If the mark (dents, scratches, etc.) cannot be transmitted, then neither can the structure, and the process cannot be causal.

But to say that a causal process transmits a mark between two points in its world line still leaves us with the question of exactly how this is accomplished. Exactly *what* does the causal relation consist of? Salmon answers this question by appealing to Russell's 'at-at' theory of motion, and I quote him at length here:

[22] When $r = 50,000km$, the circumference, $l = 2\pi r \approx 314,159km$. The speed of light is $\approx 300,000km/s$.

[23] Cf. Salmon 1984, 143-144.

> How does the mark *get from* an earlier place in the process to a later one? Putting the question in this way suggests a strong analogy. More than 2500 years ago Zeno of Elea posed the famous paradox of the flying arrow. How does the arrow *get from* point A to point B in its trajectory? ... Early in the twentieth century Bertrand Russell offered what is, I believe, a completely satisfactory resolution of the arrow paradox in terms of the so-called *at-at theory of motion*. Motion, Russell observed, is nothing more than a functional relationship between points of space and moments of time. To move *is* simply to occupy different positions in space at different moments of time. ... To get from point A to point B consists merely of being *at* the intervening points of space *at* the corresponding moments of time. ... There was, it seemed to me, a similar answer to the question about mark transmission. A mark that is imposed at point A in a process is *transmitted* to point B in that same process if, *without additional interventions*, the mark is present at each intervening stage in the process.[24].

Salmon describes precisely, and at great length, the different types of 'interventions', or interactions that may occur between processes: interactive forks, perfect forks, conjunctive forks, and so on.[25] But I will not discuss these here. For our purposes, it will suffice to say that an interaction is a joining, in spacetime, of the world lines of two or more causal processes that results in a modification of one or more of these processes. For example, the collision of the car with the wall will leave modifications, or marks, on both the car and the wall that will persist from that moment forward.[26] Salmon notes, importantly, that such modifications occur "only when (at least) two causal processes intersect. If either or both of the intersecting processes are pseudo-processes, no such mutual modification occurs."[27]

But there is a (perhaps obvious) objection due to Kitcher regarding Salmon's distinction between causal and pseudo-processes.[28] Kitcher makes the point that pseudo-processes can sometimes transmit pseudo-marks. Imagine the car again. Suppose it is involved in a fender-bender. After the accident, the car transmits the property of having a dent. But now the shadow, also, from that moment forward transmits the property of being the shadow

[24]Salmon 1989, 110

[25]See Salmon 1984, 158-183 for his detailed exposition.

[26]An intersection of this sort is a necessary, but not sufficient condition for one of the processes to be modified. It is possible for two causal processes to intersect without the intersection resulting in a modification of either, as is commonly the case with photons.

[27]Salmon 1984, 169.

[28]Kitcher 1989, 463.

of a dented car. Also, it seems as though there are many cases in which pseudo-processes actually cause modifications in causal processes. Encountering the shadow of a human being holding an axe as I walk alone down a dimly lit alleyway late at night, for example, may cause me to become afraid, and this may cause me to run away. It will not do to say that it is not actually the shadow itself that causes my fear (but the beams of photons which give it its shape, say), for here it is my *recognition* of this dark patch of ground *as the shadow* of an axe-wielding maniac which causes me to become afraid. If I later find out that it was only the shadow of the statue of Charlemagne that stands in the town square just up ahead, then my fear dissipates. Salmon's descriptions of the interactions between processes do not seem to handle the case of mental states very well.

It is difficult to say how Salmon would respond to the second objection. Perhaps he could say that mental states are always explainable in terms of their corresponding brain states. In that case it would be correct to say that it was the photon beams (among other things) and not the shadow that caused my 'fear'. But to opt for this strategy is to take a very reductionistic view of cognition. I think this is undesirable and unnecessary for a theory of causality, let alone for a theory of scientific explanation, but let us leave this issue aside for now. I will have a great deal more to say about such matters in the pages that follow. With regards to the first (Kitcher's) objection, what I think that Salmon should say is that although the shadow does, in a sense, 'transmit' the property of being a dented shadow, this is not the same sense of 'transmit' that he is using; for it is obviously not the case that this property is being transmitted *by the shadow* along its world-line. Rather, for the shadow to continually display this property, it requires that the car be continuously dented. The fact that the shadow exhibits a dent at one moment is not the cause of its exhibiting a dent at subsequent moments. It is obviously not *the shadow* that is transmitting this property. However such a response cannot be made by Salmon using the resources of the at-at theory alone, for according to the at-at theory, it makes perfect sense to say that the shadow is transmitting a mark, for *according to all appearances*, this is exactly what it is doing, and at-at only gives us a rule that can be applied to appearances.

But the failings of the mark transmission theory aside, one thing that we can say in Salmon's defence is that there obviously are fundamental differences, intuitively at least, between causal and pseudo-processes. Salmon's examples clearly demonstrate this (who would deny that there is something fundamentally different about a shadow as opposed to a car, or about an illuminated spot on a wall as opposed to an actual beam of photons?). Perhaps mark transmission is too naïve a model, but that does not mean that Salmon's basic idea is wrong-headed, for perhaps there is some other criterion that we can identify that can distinguish between causal and pseudo-processes.

One solution (suggested by van Fraassen) is to say that our best scientific theories *imply* that some processes cannot transmit marks; therefore *relative to our theories*, some processes are causal and some processes are not.[29] In fact, in his later writings, Salmon abandons the mark transmission criterion altogether and formulates a theory of causal processes in just the way suggested. Heavily influenced by Phil Dowe, Salmon eventually comes to accept a 'conserved quantity' criterion.[30]

Salmon (i.e., his later view) and Dowe actually agree in a great many respects. However, Dowe, unlike Salmon, foregoes the requirement of transmission. On Dowe's view, causal processes and causal interactions are defined as follows:

> CQ_1. A *causal interaction* is an intersection of world lines which involves exchange of a conserved quantity.
>
> CQ_2. A *causal process* is a world line of an object which possesses a conserved quantity.[31]

Conserved quantities are simply those quantities that are universally conserved according to current scientific theory: mass-energy, linear momentum, and so on. Shadows count

[29] van Fraassen 1980, 121.

[30] In Salmon, 1994, in response to the criticisms of Dowe and others, he abandoned the mark transmission criterion in favour of an 'invariant quantity' criterion, only to later retract that as well and accept Dowe's 'conserved quantity' criterion, in Salmon, 1997.

[31] Dowe 1995, 323.

as objects, according to Dowe, just as much as cars do; however they are not the right kind of object in the sense that they are incapable of possessing the right type of conserved quantity (those quantities which are universally conserved according to theory). A shadow has shape, position, velocity, and so on; but it cannot possess momentum or energy.

For Dowe, the main work done by Salmon's transmission criterion is to give causal processes a sense of directionality. The conserved quantity, on Salmon's view, must be transmitted *from* point A *to* point B (note that this does not, in principle, rule out backwards-in-time causation). But according to the at-at theory, transmission from A to B amounts to the possession of the conserved quantity at every spacetime point along the interval. Dowe points out that "the left hand side of this definition has directionality ("from A to B") but the right hand side does not ("possession at each spacetime point"). This means that transmission, once defined, turns out to lack directionality. Thus the account appears to offer an account of causal direction but in fact fails to deliver the goods."[32] Given this, the requirement of transmission becomes superfluous, for there is nothing involved in it that is not already captured by Dowe's notion of the possession of a conserved quantity along the world line of an object.

But defending the transmission requirement, Salmon asks us to consider the example of the spotlamp and the spot again. If we consider the world line, not of the spot itself, but of *the portion of the wall illuminated* by the spot, then it is plain that this object possesses a conserved quantity (the energy it receives as a result of its interaction with the photons in the beam) all along its world line. Now we would be loath to call this object a causal process, of course, and the reason why, according to Salmon, is that this object does not transmit any energy; its energy is being received at all times from an exterior source.[33]

But for Dowe, Salmon's objection is misguided. He responds that the 'object' that Salmon is referring to (the part of the wall illuminated by the spot) is not actually an object at all; it is, rather, merely a *time-wise gerrymandered aggregate*, which amounts to nothing

[32] Dowe 1995, 326.
[33] Salmon 1994, 308.

more than 'spatiotemporal junk'; and whatever is not an object cannot be a causal process either. For example. Consider an 'object', x, which we define as:

for $t_1 \leq t < t_2$	x is the coin in my pocket.
for $t_2 \leq t < t_3$	x is the red pen on my desk.
for $t_3 \leq t < t_4$	x is my watch.[34]

Here, x, occupies a determinate region in spacetime, and for $t_1, ..., t_4$, x possesses conserved quantities. Another time-wise gerrymander is "the president of the United States"; yet another is "the object nearest to my car", and yet another: "the molecule in this box with momentum p". One more is "the part of the wall illuminated by the spot as it moves about the stadium". Now do the world-lines of x and of these other time-wise gerrymanders qualify as causal processes?

> The answer is no, because x does not qualify as an object. ... I take it that *an object must be wholly present at a time* in order to exist at that time. That is, when you have an object at a time, it is not that strictly speaking you just have a part of that object - a temporal part. If you have an object then you have the whole object at that time. In other words, time-wise gerrymanders are not objects.[35]

As an alternative to defining an object as something that is 'wholly present at a time', Dowe considers the option of taking objects to be four-dimensional; i.e., of considering an object as existing in time in the same way that it exists in space, and thus as having both temporal and spatial parts. But doing so makes individuation vastly more complicated: "Not only do you have to worry about defining the object at a time, but also over time. Which four-dimensional objects are genuine objects, and which are not? Which splotches

[34]Dowe 1995, 328.
[35]Dowe 1995, 329.

on the space-time diagram are really objects?"[36] Dowe also makes the point that this characterization of an object (as something existing *in* time) is one that is built into our scientific theories:

> Scientific quantities are ascribed to objects in a way that is essentially temporally "localized," ... Scientific quantitative properties are possessed at a time, where the instant is thought of as a point. There is no spatial analogy: sometimes planets or particles are in some respects treated as points in space, possessing, for example, momentum, but this is always recognized as an idealization. ... But in the case of a putative object consisting of a time-wise gerrymander of genuine objects with momentum, there is no sense in which for example the sum of the quantities can be ascribed to some object ... So, given that we are trying to distinguish genuine from pseudo objects via possession of scientific quantities, it makes sense to adopt the concept of an object as wholly present at a time.[37]

But with respect to Dowe's example of the 'object' consisting of the coin, the red pen, and the watch, Salmon objects that this is just a red herring: on Salmon's view, causal processes are spatiotemporally *continuous* (i.e., they have a continuous world line), and unlike the part of the wall illuminated by the spot, x does not count as an object on his view either (although each piece of it does). With respect to the illuminated portion of the wall, Salmon makes the point that the concept of genidentity (identity over time), which Dowe uses but does not analyze, is just as difficult a concept whether we take objects to be three-dimensional or four-dimensional. We are probably all aware of the famous philosophical problem of the ship of Theseus;[38] or consider the following more abstract example: suppose that a completed jigsaw puzzle comes into existence at time t_0. If at time t_1 we take the puzzle apart, then it goes out of existence. But if at time t_2 we put the pieces back together again, is the completed puzzle at time t_0 the same or different from the completed puzzle at time t_2? Things are particularly difficult in the quantum domain, where in some cases it is impossible *in principle* to determine genidentity, for example when two identical particles

[36]Dowe 1995, 330.
[37]Dowe 1995. 330-331.
[38]The ship on display in ancient Athens that over time, had all of its original parts replaced.

collide.[39]

But in light of certain aspects of Salmon's own theory, Salmon's criticisms of Dowe strike me as strange. Salmon's view, if we take it literally, seems to entail that a baseball or a bullet flying through the air cannot actually count as a single causal process. After explaining his invariant quantity theory (the second incarnation of his transmission view), he writes:

> Speaking literally, the foregoing definitions imply that a causal process does not enter into any causal interactions. For example, a gas molecule constitutes a causal process between its collisions with other molecules or the walls of the container. When it collides with another molecule, it becomes another causal process which endures until the next collision. ... When we understand this technical detail, there is no harm in referring to the history of a molecule over a considerable period of time as a single (composite) causal process that enters into many interactions.[40]

With regard to bullets or baseballs flying through the air (and interacting with atmospheric molecules), he writes:

> In this, and many similar sorts of situations, we would simply ignore such interactions because the energy-momentum exchanges are too small to matter. Pragmatic considerations determine whether a given "process" is to be regarded as a single process or a complex network of processes and interactions.[41]

Salmon cannot have it both ways. He cannot say that for pragmatic reasons we should forget about the interactions between the baseball and the air molecules - that we can idealize this process for the purposes of explaining the baseball's flight - while at the same time deny to Dowe that he can treat Theseus' ship, as it was in the beginning, and Theseus' ship as it was after all of its original parts had been replaced, as the same object - in other words that he can idealize the ship for pragmatic reasons. Just as we can decide, for pragmatic

[39]Cf. Feynman 1965 §§3-4.
[40]Salmon 1994, 308-309.
[41]Salmon 1994, 309.

reasons, what we will count as a process, we can decide, again for pragmatic reasons, what we will count as an object existing wholly in time from one moment to the next.

Further, Dowe's second point has not been addressed. It is still the case that scientific quantities are ascribed, in our theories, to objects situated in time. While there is nothing barring us from characterizing objects in four dimensions, it is more convenient, and it simply makes more sense, to be consistent with the language of our scientific theories in our description of causal processes. Perhaps it is even necessary. This last point is one that I will return to in my final chapter.

But whatever the merits of the accounts of causality that we have considered here are, as accounts of explanation *in general* they are, if not deficient then at least incomplete; for they say nothing about explanations which *do not* involve causes at all; and these certainly do exist. There are explanations in terms of structural laws, such as the Pauli exclusion principle[42] and Boyle's law for gases.[43] Kitcher points to the cases of formal linguistics and mathematics:

> Two obvious examples are formal syntax and pure mathematics. Explanations of the grammaticality or ungrammaticality of particular strings in particular natural languages are given by identifying the constraints set by the underlying rules of syntax. Note that it will not do to suggest that the formal explanation is a placeholder for a description of causal processes that occur in the brains of speakers - for part of the point of the enterprise is to distinguish between explanations of competence and explanations of performance.[44]

Later, I will examine one decidedly non-causal class of explanations: those given according to the Copenhagen interpretation of quantum mechanics. However with respect to the examples above, it is, I think, possible to contest them (although I do not). It might,

[42]The Pauli exclusion principle is a purely formal law to the effect that no two identical fermions (particles with a half-integer spin, e.g., protons, electrons, etc.) can simultaneously occupy the same quantum state. This explains why, in an atom, only two electrons are allowed per orbit. Since with respect to a particular orbit there are only two possible spin values for the electron, only two electrons (with opposite spin values) are permitted to occupy the same orbit at any one time.

[43]Which states that the temperature of a gas is proportional, at any given time, to the volume times the pressure.

[44]Kitcher 1989, 423.

for example, seem odd, for some, to call a mathematical derivation an *explanation.* As for the structural laws, perhaps they can in principle be further explicated in terms of causal mechanisms? But whatever the case may be, an account of explanation should not force us to commit ourselves in either direction. An ideal account of explanation should, if possible, be capable of handling both causal and non-causal cases.

1.3 Pragmatic Aspects of Explanation

A second aspect of explanation that is neglected by Hempel's D-N model is the so-called pragmatic aspect. If we consider an explanation to be an answer to a 'why-question' ('why is the sky blue?', 'why is my brother so mean?', etc.), then the answer to a why-question will depend on the context in which it is asked, and what in certain contexts may seem like an inappropriate explanation will in other contexts constitute a perfectly good one. Consider the following variation on the 'flagpole' example, in which the length of a tower's shadow explains its height:

> That tower marks the spot where he killed the maid with whom he had been in love to the point of madness. And the height of the tower? He vowed that shadow would cover the terrace where he first proclaimed his love, with every setting sun - that is why the tower had to be so high.[45]

For van Fraassen, all explanations are answers to why-questions; questions of the form "Why (is it the case that) P *in contrast to* (other members of) X?", where the second half of the question is taken as implicit in context and typically left unstated. X is called the contrast class, and is a set of alternatives to P. For example, the question: "Why did you dye your hair black?" could be interpreted as either "Why did you dye your hair black, as opposed to blond or blue or orange?", or alternatively, "Why did you dye your hair black, as opposed to not dying it at all?". An answer to a why-question will be one that *favours* P over any of its alternatives in the given contrast class.

[45]van Fraassen, 1980, 133-134.

Now imagine I ask you the question, "Can you get to Victoria both by ferry and by plane?" And consider the following alternative answers:[46]

(a) Yes.

(b) You can get to Victoria both by ferry and by plane.

(c) You can get to Victoria by ferry

(d) You can get to Victoria both by ferry and by plane, but the ferry ride is not to be missed.

(e) You can certainly get to Victoria by ferry, and that is something not to be missed.

(b), here, is called a *direct answer.* It provides just enough information to answer the question. (a) is shorthand for (b). (c) and (e) are partial answers, for they are implied by the direct answer. (d), finally, is called a complete answer, for though it supplies additional information, it implies the direct answer. We next define the following concepts:

> a *presupposition* of question Q is any proposition which is implied by all direct answers to Q.
>
> a *correction* (or *corrective answer*) to Q is any denial of any presupposition of Q.
>
> the (*basic*) *presupposition* of Q is the proposition which is true if and only if some direct answer to Q is true.[47]

A question, Q, arises in a given context, when given that context, a certain body of background information, K, implies the central presupposition of Q. If the central presupposition of Q is not implied by K, our answer is simply that the questioner is mistaken. For example, suppose I ask you the question, "Why does Santa Claus live at the north pole?" In this case, the background knowledge, K, which is accessible to both of us, implies that there is no Santa Claus. Therefore I retort with the *corrective* answer: "But you are mistaken. Santa Claus does not exist."

[46]Cf. van Fraassen 1980, 138.
[47]van Fraassen 1980, 140.

Now consider the following question, "Why is this engineer insane?". We can interpret this more formally as Q = "Why is it the case that this engineer is insane in contrast to those other engineers?". Now P is the proposition that this engineer is insane. We call it the *topic of concern.* It is also the basic presupposition of Q. The contrast class, X, is "those other engineers". Finally, there is what van Fraassen calls the 'respect-in-which a reason is requested': the relevance relation, R. An answer, A, counts as being explanatorily relevant to a question if R holds between A and $< P_k, X >$. Relevant answers might consist, for example, of certain events in this engineer's childhood, a strained relationship with her co-workers, etc. It is important to note that we cannot judge the relevance relation in isolation; "... we must say of a given proposition that it is or is not relevant (in this context) to the topic with respect to that contrast class."[48]

Thus we can define a why-question, Q, according to the following schema:

$$Q =< P_k, X, R >$$

Where P_k is the topic, X is the contrast class, and R is the relevance relation. And an answer to Q will be expressed as follows: P_k *in contrast to* (the rest of) X *because* A. We can now define a *direct answer* in the following way:

> B is a *direct answer* to question $Q =< P_k, X, R >$ exactly if there is some proposition A such that A bears relation R to $< P_k, X >$ and B is the proposition which is true exactly if (P_k; *and* for all $i \neq k$, *not* P_i; and A) is true.[49]

Van Fraassen's theory of the *pragmatics* of explanation is actually put forward by him as an overall theory *of explanation*; however taken in the latter sense it runs into problems on account of the fact that he imposes absolutely no restrictions on the relevance relation. As Salmon writes:

[48] van Fraassen 1980, 142.
[49] van Fraassen 1980, 144.

> Suppose that someone asks why John F. Kennedy died on 22 November 1963; this is the question $Q = < P_k, X, R >$, where
>
> P_k = JFK died 11/22/63
>
> X = {JFK died 1/1/63, JFK died 1/2/63, ... JFK survived 1963}
>
> R = astral influence
>
> Suppose that the direct answer is P_k in contrast to the rest of X because A, where A ... consists of a *true* description of the configuration of the planets, sun, moon, and stars at the time of Kennedy's birth. Suppose further that the person who supplies the answer has an astrological theory from which it follows that, given A, it was certain, or highly probable, that Kennedy would die on that day. We now have a why-question and an answer; the answer is an explanation. We must ask how good it is.[50]

Second, as impressive as van Fraassen's theory of 'why-questions' is, it ignores other types of questions, so-called 'how-possibly' questions. For example, 'How can Santa Claus possibly manage to deliver all those toys in just one evening?'. This question does not ask for the reason why Santa Claus does this, but for a description of how he does it. It implies that the answer should consist of an account of the special characteristics of the sled and of the reindeer, the circumference of the earth, how many deliveries need to be made, and so on. Of course, this how-possibly question could be translated into something like: 'Why is it that Santa Claus is able to deliver all those toys in just one evening?'. But I do not think this will do. I can always answer that why-question with something like: 'because he can afford to buy the proper equipment'. The how-possibly question, in contrast, seems to demand that we explain exactly how it is that his equipment is 'proper'. Now I do not want to venture further into this debate, for I think that the theory of explanation which I am about to introduce can be seen as an 'umbrella' theory under which both the causal and pragmatic aspects of explanation can fit, minus these objectionable consequences. This is Peter Railton's theory, and it is the subject of the next section.

[50] Salmon 1989, 142. See also: Kitcher & Salmon, 1987.

1.4 The Deductive-Nomological-Probabilistic Model

The view that, to my mind, best deals with both pragmatic and causal concerns is that of Peter Railton. Railton accepts the D-N model as a starting point. However, unlike Hempel, whose conception of explanation is epistemic, Railton's conception (like Salmon's) is *ontic*. For Railton, the goal of explanation is not to give reasons to expect the occurrence of a phenomenon, but to give an account of the actual *mechanisms* that underlie the phenomenon:

> While prediction and control may exhaust our practical problems in the natural world ... explanation is an activity not wholly practical in purpose. The goal of understanding the world is a theoretical goal, and if the world is a machine - a vast arrangement of nomic connections - then our theory ought to give us some insight into the structure and workings of the mechanism, above and beyond the capability of predicting and controlling its outcomes ... Is the deductive-nomological model of explanation therefore unacceptable? - No, just incomplete. Calling for an account of the mechanism leaves open the nature of that account, and as far as I can see, the model explanations offered in scientific texts are D-N when complete, D-N sketches when not.[51]

Railton purposely leaves the definition of mechanism vague; but this is a virtue of his model, for we can leave it open whether, for example, Salmon's or Dowe's characterization of causal processes is the correct one. Further, his 'account of the mechanism' need not appeal to causes at all where these are inapplicable (e.g., in the case of the Pauli exclusion principle that I mentioned above, and in general for what he calls 'structural laws'[52]). The requirement to give an account of the mechanism can thus be read as a requirement to have our explanations be as deep and as detailed (and as true) as science allows. On this view, the D-N explanation of the flagpole's height in terms of its shadow (but not van Fraassen's version of it) is clearly unsatisfactory: it tells us nothing about the (in this case, causal) processes from which the flagpole resulted. A good explanation would have to

[51]Railton 1978, 208.

[52]These include, e.g., the second law of thermodynamics, the principle that the speed of light is constant in all inertial frames, and so on.

involve an appeal to the materials from which the flagpole was constructed, the method of its construction, perhaps a description of its intended use, and so on.

The requirement to give an account of the mechanism also allows us to satisfactorily answer a particularly difficult counter-example to Hempel's version of the D-N model.[53] This is Peter Achinstein's 'intervening-cause' counter-example, where what in normal circumstances constitutes a perfectly good explanation for an occurrence is made irrelevant by some intervening event. For example, suppose someone, call him John Jones, eats a pound of arsenic at time t. Call this circumstance C_1, and let L_1 be the law such that anyone who eats a pound of arsenic at time t will be dead within 24 hours of t. Now if Jones is dead after only 3 hours (the explanandum event, E), then in that case, C_1, L_1, and E make up a perfectly good D-N argument for Jones' death (this argument satisfies all of Hempel's criteria and we can infer E from C_1 and L_1). But suppose that Jones is actually hit by a bus at time $t + 3h$. That Jones ate a pound of arsenic is irrelevant to his actual death - for the *bus* caused his death. But our D-N argument fails to capture this. On Railton's account, however, any explanation that does not completely trace the (in this case, causal) history of the events leading up to his death would be insufficient - we would be *required* to mention the bus (as well as the fact that Jones was found bruised and bleeding underneath it).

For Railton, D-N explanations are actually a special case of what he calls the Deductive-Nomological-Probabilistic (D-N-P) model. In its complete form, the D-N-P is a schema through which we can explain the occurrence of any phenomena, whether that occurrence is deterministic or indeterministic, probable or improbable. It is represented (my paraphrase) as follows:

(1) A *complete* theoretical account of the mechanism involved in the production of the phenomenon described by the explanandum statement, and a derivation, from our theory, of the laws required to infer the probability of occurrence of the explanandum phenomenon, E.

(2) A D-N argument from the laws derived in (1) to the probability of occurrence of E.

[53] Cf. Achinstein, 168.

(3) A statement to the effect that E actually did occur. [54]

The D-N argument, while still a necessary condition for explanation, is assigned a relatively small role. Most of the work is actually done in step (1). Note, also, that what we infer from the D-N argument is not the occurrence of the explanandum event, but merely the probability of its occurrence. The last step, what Railton calls the 'parenthetic addendum', is used to chain one or more explanations together. It is actually required only for probabilistic explanations (which I will discuss in more detail in the next chapter). In the case of deterministic phenomena, the parenthetic addendum is redundant and can be left out (since the probability of its occurrence is 1).

Step (1), as we saw, requires us to give a *complete* theoretical account of the laws and conditions invoked in step (2). But one might object that Railton's requirement places an extremely heavy burden on the explainer (and on the person being explained to). Indeed, it seems likely that an exhaustive search through the entire class of introductory level (likely even advanced) science textbooks would fail to turn up even one explanation that comes close to meeting Railton's requirement. Moreover, the laws that we give an account of, for Railton, must be *true* laws, in a very strict sense. Strictly speaking, we cannot appeal to statistical laws in our explanans where these serve merely as approximations of the behaviour of deterministic phenomena. Thus, statistical laws that explain weather patterns, or roulette wheels - indeed most of the statistical laws that are actually used in every day scientific practice - cannot figure in an ideal explanation. Railton does not see this as problematic:

> The use of epistemic or statistical probabilities in connection with such phenomena unquestionably has instrumental value, and should not be given up. What must be given up is the idea that *explanations* can be based on probabilities that have no role in bringing the world's explananda about, but serve only to describe deterministic phenomena.[55]

[54] Cf. Railton 1978, 214, 218.
[55] *Ibid.*, 223.

Railton's exhortations aside, one might still feel a little uneasy about 'giving up' as 'not an explanation' so much of the explaining that scientists actually do. A related concern is what to say about explanations involving the laws, e.g., of Newtonian physics. On Hempel's view, recall, the use of these derivative laws is unproblematic. But it is not clear whether something similar can be done - in practice - on Railton's account, for if the requirement is to give as deep and as detailed an explanation as possible, then even if, technically, we do not rule out derivative laws, it does not appear that we are ever in a position to use them; for it is hard to see how any explanation of, e.g., the motion of the planets, can be given without appealing to the laws and principles of general relativity.

Railton's answer to these concerns is his notion of 'explanatory information'. He writes:

> Whatever the reason, does the D-N-P account force us to say that these less-than-full-fledged specimens are not explanations?
>
> In many cases, that would be an intolerably strict position to take. But where should one draw the line between explanation and non-explanation? The answer lies in not drawing lines, at least at this point, and in recognizing a continuum of explanatoriness. ... At one end we find what I will call an *ideal D-N-P text* ... At the other end we find statements completely devoid of what I will call 'explanatory information'. What is explanatory information? Consider an ideal D-N-P text for the explanation of a fact p. Now consider any statement S that, were we ignorant of this text but conversant with the language and concepts employed in it and in S, would enable us to answer questions about this text in such a way as to eliminate some degree of uncertainty about what is contained in it. To the extent that S enables us to give accurate answers to such questions, i.e., to the extent that it enables us to reconstruct this text or otherwise illuminates the features of this text, we may say that *S provides explanatory information concerning why p*.[56]

Thus we do not need to rewrite all of our science textbooks; they succeed in conveying explanatory information. The case is similar for Newtonian physics. As for the (not irreducibly) statistical laws: "... it should be no mystery that statistical generalizations believed *not* to reflect underlying probabilistic laws may still be explanatory, once we recognize that

[56]Railton 1981, 240.

they function in explanation not as ersatz laws but as summaries of information about initial conditions and boundary conditions"[57]

A further advantage of the notion of explanatory information is that it brings pragmatic considerations into the picture. If I am confronted with a a why-question, I can convey whatever part of the explanatory text I consider to be salient to the questioner (this will depend on the knowledge situation, K that both I and the questioner share, and also on the exact why-question, $Q =< P_k, X, R >$ that I am asked). And since the relevance relation is *constrained by the explanatory text*, whatever answer is given must correspond to something *in that text*. There is no possibility of giving a proper explanation of JFK's death in terms of the configuration of the planets at the time of his birth. Thus while salience is a subjective matter, relevance is an objective one.[58]

But now even with the notion of explanatory information, Railton's account of explanation may still make us feel uneasy, for consider his comments on the ideal explanatory text:

> ... an ideal text for the explanation of the outcome of a causal process would look something like this: an inter-connected series of law-based accounts of all the nodes and links in the causal network culminating in the explanandum, complete with a fully detailed description of the causal mechanisms involved and theoretical derivations of all the covering laws involved. This full-blown causal account would extend, via various relations of reduction and supervenience, to all levels of analysis, i.e., the ideal text would be closed under relations of causal dependence, reduction, and supervenience. It would be the whole story concerning why the explanandum occurred, relative to a correct theory of the lawful dependencies of the world. Such an ideal causal D-N text would be infinite if time were without beginning or infinitely divisible, and plainly there is no question of ever setting such an ideal text down on paper.[59]

Thus it appears that an (ideal) explanation of any phenomenon must go all the way down to the level of quantum physics if it is to be complete. Now we are not *required* to do so in order to convey explanatory information:

[57] *Ibid.*, 251.
[58] Salmon also makes this point, in: Salmon 1989, 162.
[59] Railton 1981, 247.

> ... within the division of labour among scientists it is possible to find [some group] interested in developing the ability to fill in virtually any particular aspect of ideal texts ... A chemist may be uninterested in how the reagents he handles came into being; a cosmologist may be interested in just that; a geologist may be interested in how those substances came to be distributed over the surface of the earth; an evolutionary biologist may be interested in how chemists (and the rest of us) came into being ... [60]

Still, I am not sure that, e.g., conveying explanatory information about the radioactive properties of the atoms in Jones' brain tells us *anything* about, say, Jones' barophobia.[61] I, for one, do not believe that (hard) reductionism is true. This is, of course, arguable. But whether or not hard reductionism is true, I do not believe that an account of scientific explanation should force us to commit to it. Thus while it may seem like nit-picking on my part, I suggest that we weaken Railton's account. Instead of maintaining that there be one ideal explanatory text for a given fact, why not say instead that there can be many (e.g., one per special science)? Whether or not any one of these texts is reducible to any other then becomes a separate question. Let the debate rage on. As long as our account of explanation keeps these texts separate, the outcome of that debate cannot impact us.

Now one might object along the following lines: "even if I give you that you may have more than one ideal text for a fact, are they not still ideal texts for *one and the same phenomenon*? How is your position any different from saying, as Railton does, that all of these represent different *aspects* of the same phenomenon, and that therefore they represent different *aspects* of one and the same ideal explanatory text for that phenomenon?" My point, however, is that they are *not* the same phenomenon; "barophobia" is *not* a phenomenon that one encounters in quantum physics. "Alpha-decay" is *not* a phenomenon that one encounters in psychology; for the *concepts* that we use to describe phenomena in these two sciences are not, to my mind, translatable. The fact that these phenomena happen to be associated with the same spacetime point (which is yet another conceptualization) does not entail that they are identical phenomena. Again, this is debatable. I will come back to this

---

[60] *Ibid.*

[61] The fear of gravity.

issue in chapter 3, but for now, let me just say that it would *prima facie* be useful, from the standpoint of an account of explanation, to at least allow that it is possible that the special sciences are not all reducible to physics. If our account of explanation stands or falls on this issue, so much the worse for us.

We have come to the end of this chapter. Before we conclude, let us take the opportunity to look back at the ground we have covered so far. We discussed four models of explanation: Hempel and Oppenheim's D-N model, according to which explanations are arguments that proceed deductively from a set of statements describing laws and initial conditions, to a statement describing the fact to be explained; Salmon's C-M model, according to which an explanation is a detailed tracing out of the causal history of an event; van Fraassen's theory of why-questions; and finally, Railton's D-N-P model which attempts to retain the best aspects of all three of these.

With respect to Hempel's D-N model, one of its virtues is that it is metaphysically neutral. It takes no stance on whether the laws appealed to in explanations are literally true, and it takes no stance on the issue of whether some, or all, or none of the explanations given in scientific contexts are causal explanations. On the other hand, I argued that the D-N model fails to account for the asymmetrical character of causal explanations, and that it fails to account for the role that pragmatic considerations often play - for the fact that the worth of an explanation is usually judged with respect to the context in which it is given. With respect to the causal models of explanation, we saw two accounts of causality in terms of causal processes: Salmon's, for which the emphasis is on transmission (either of a mark or of a conserved quantity), and Dowe's, for which the emphasis is on the idea of an object, existing in time, endowed with causal powers (i.e. with conserved quantities).[62] I argued that purely causal models, though promising as accounts of causation, are deficient as accounts of explanation in general because they fail to account for the many non-causal explanations that actually occur in scientific contexts. Van Fraassen's pragmatic model

[62]This notion of an object as something existing wholly at a time is actually quite important and it is something that I will return to in my third chapter.

(his theory of why-questions and their answers) likewise is promising as an account of the *pragmatics* of explanation, however as an account of *explanation* it is deficient due to the fact that it does not place any restrictions on the relevance relation between an answer and the question asked. In sum: the pragmatic theory does not restrict permissible explanations enough, while the causal theory restricts them too much.

Railton's account is in a certain sense a compromise position. Railton's D-N-P model keeps the D-N model as a necessary, but not sufficient condition for explanation; the D-N-P's requirement to give a thorough account of the mechanism (while leaving the actual nature of the mechanism open) is sufficient to satisfy the causal theorists; Railton's notion of explanatory information is sufficient to satisfy those, such as van Fraassen, who place more emphasis on the pragmatic aspects of explanation; and by combining these it seems as though he has eliminated the main problems with all of them. But I argued that Railton has not gone far enough. For his ontic account of explanation still makes too many metaphysical assumptions. We have yet to see all of these, but what we have seen so far is that Railton's view entails that there can be only *one true ideal* explanation for any phenomenon. I argued, however, that we should reject this requirement and that we should not commit ourselves to such an extreme reductionistic picture of science.

It might seem, however, as though I have robbed the D-N-P of its heart and soul. It is the principal requirement of Railton's view, recall, to go as deep as possible in our explanations in order to get at the *truth*. But if we eliminate this requirement, then what is left? It looks, now, as though the D-N-P is no better at handling the problem of the relevance relation than van Fraassen's theory, and is not the door wide open to just any old explanation, astrological or otherwise? No. For I have not argued that we should forego the *goal* of explaining phenomena as deeply and as truly as possible. And I have not denied that an ideal explanation is the one that is maximally deep and maximally true. I have simply made the point that in giving an account of explanation we should allow that sometimes it may (perhaps) be *impossible in principle* to reduce one level of description to another. Now this may be at odds with Railton's own views on the subject. But I am not defending Railton here, only his model.

Chapter 2

Explanations of Indeterministic Phenomena

We now leave deterministic phenomena and begin our discussion of explanations of indeterministic phenomena. The two views that I will focus on here are Hempel's Inductive-Statistical (I-S) account of explanation and Railton's D-N-P account of probabilistic explanation (which I introduced in the last chapter). But before we begin our discussion of probabilistic explanation, we must first discuss the concept of probability itself and of its various interpretations. I will focus, here on the interpretations that are most relevant to our purposes: the classical, frequentist, logical, and propensity interpretations, as well as Carnap's hybrid view.

2.1 Probability and its Interpretations

When the outcome of a physical process is uncertain, we say that that outcome has a particular probability to occur. For example, on any given spin of a roulette wheel, it is practically impossible (though perhaps possible in principle) to describe the initial conditions of the spin precisely enough to determine exactly which square the ball will land on. To make such a determination, we would need, for instance, to take account of the velocity of the wheel, the motion of the croupier's hand, the temperature and humidity of the air in the room, the material out of which the wheel and ball are made, and a host of other conditions. Instead, we say that for any given spin, the probability of, say, landing on a red space is .4737 (i.e., 47.37%).[1] But what is the *meaning* of a probability statement? Does it assert an objective characteristic about the physical process we are describing, or is it a measure of our knowledge of the situation under consideration? This is the question to which I now turn.

The Probability Calculus

Before we begin, however, it will be useful to briefly describe the basic mathematical properties of probability statements, defined by the so-called 'probability calculus'. The probability calculus, as defined by Kolmogorov, rests on three basic axioms (the Kolmogorov axioms), which we can illustrate informally as follows. Consider the set of possible results for the roll of an ordinary die: $\{1, 2, 3, 4, 5, 6\}$. Call this set Ω. Let F be the set of all non-empty subsets of Ω, i.e. $\{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1, 2\}, \{1, 3\}, \ldots \{1, 2, 3\}, \{1, 2, 4\}, \ldots \{1, 2, 3, 4, 5, 6\}\}$. Now let A be some arbitrary element of F. We then call $P(A)$ the 'probability of A'. For example, if $A = \{2, 4, 5\}$, then $P(A)$ is the probability that a die lands on either 2 or 4 or 5. According to the first axiom of the calculus, $P(A)$ is always a non-negative real number, i.e.:

$$(1) \quad P(A) \geq 0, \quad \forall A \in F$$

Now recall that Ω represents all possible results. Thus $P(\Omega)$, in this case, will be the probability of landing on either 1 or 2 or 3 or 4 or 5 or 6. Since these exhaust all possibilities, we can say that it is certain that one of the outcomes belonging to Ω will

[1] In American roulette.

obtain. According to the second axiom, we represent certainty by the value 1. Thus,

$$(2) \quad P(\Omega) = 1$$

Finally, the probability that either A or B occurs (where A and B are subsets of F and have no elements in common) is given by the addition rule, which is the third axiom:

$$(3) \quad P(A \cup B) = P(A) + P(B), \quad \forall A, B \in F, \text{ where } A \cap B = \emptyset$$

For example, $P(\{1\} \cup \{2\}) = P\{1\} + P\{2\} = 1/6 + 1/6 = 2/6$. We can derive all the other theorems of the calculus from these three axioms. For example, the multiplication rule, for the probability that *both* A and B occur (where A and B are independent events[2]) is given by:

$$P(A \cdot B) = P(A) \times P(B)$$

One important concept is that of *conditional* probability, the probability of A given B. For example, in scientific contexts we often speak of the probability of an hypothesis given the evidence. We can define this in terms of the multiplication rule as follows:[3]

$$P(A|B) = P(A \cdot B)/P(B)$$

While the probability calculus gives us rules for computing complex probabilities from prior probabilities, it says nothing about how the prior probabilities themselves (e.g., $P(A)$) should be determined. In other words, the question of the *meaning* of primitive probability statements is left open. But while the probability calculus does not define the meaning, it does place constraints on it; for any interpretation of probability must satisfy the rules of the calculus.

[2]Imagine a roll of two dice, where A refers to the outcome of the first and B to the outcome of the second.
[3]In some books, the symbols are reversed; i.e., 'the probability of A given B' is written as $P(B|A)$

The Classical Interpretation

What we now call the classical interpretation of probability was formulated by Pierre Simon Laplace in the early part of the 19th century. On Laplace's view, the world is a vast causal network, thoroughly deterministic in every respect. If it were possible for us to have the eye of a deity; i.e., a comprehensive knowledge of the universe, of its every system and of its every state, we would have no need for statements of probability; for everything would be provable through the application of deterministic laws. As we are equipped with finite minds, however, all that we can hope for is an approximation of this ideal, which we approach by means of probability statements.

The basic principle behind Laplace's theory is the 'Principle of Insufficient Reason'.[4] According to this principle, if I have no reason to prefer one outcome to any other, I must consider all of them as equally possible.

> La probabilité est relative en partie à cette ignorance, en partie à nos connaissances. Nous savons que sur trois ou un plus grand nombre d'évènemens, un seul doit arriver; mais rien ne porte à croire que l'un d'eux arrivera plutôt que les autres. Dans cet état d'indécision, il nous est impossible de prononcer avec certitude sur leur arrivée. Il est cependant probable qu'un de ces évènemens pris à volonté, n'arrivera pas; parce que nous voyons plusiers cas également possibles qui excluent son existence, tandis qu'un seul la favorise.[5]

For example, consider a six-sided die. Assuming the die is not biased, I cannot say that rolling a 3 is any more likely than rolling another number, for as far as I can tell, all of these outcomes are equally possible. But now once we have divided the possible outcomes of an event into a number of equally possible cases, the probability of a particular outcome becomes a simple ratio: of the cases that are 'favourable' to an outcome, to the total possible cases:

[4]This is sometimes called the 'Principle of Indifference'.
[5]Laplace 1825, 7.

> La théorie des hasards consiste à réduire tous les évènemens du même genre, à un certain nombre de cas également possibles, c'est-à dire, tels que nous soyons également indécis sur leur existence; et à déterminer le nombre de cas favorables à lévènement dont on cherche la probabilité. Le rapport de ce nombre à celui de tous les cas possibles, est la mesure de cette probabilité qui n'est ainsi qu'une fraction dont le numérateur est le nombre des cas favorables, et dont le dénominateur est le nombre de tous les cas possibles.[6]

For example, if I want to calculate the probability of rolling a 2, the number of cases favourable to this outcome is one. If, on the other hand, I would like to calculate the probability of rolling an even number, the number of cases favourable to this outcome is three (i.e., 2, 4, and 6). To get the probability, I divide the cases favourable to the outcome by the total number of possible cases. The probability of rolling a 2, therefore, is 1/6, and the probability of rolling an even number is 3/6 = 1/2.

The Frequency Interpretations of von Mises and Reichenbach

The classical interpretation was the dominant interpretation of probability for many decades, and it accords very well with our intuitions concerning games of chance (e.g., craps, roulette wheels, lotteries, etc.). However applying it more generally, as we will see shortly, proves to be problematic. Thus by the mid-19th century, an alternative interpretation began to emerge largely through the work of Leslie Ellis and John Venn. This came to be called the *frequency* interpretation, and its first rigorous formulation was given by Richard von Mises early in the 20th century.

To help to understand the motivation behind the frequency interpretation, consider the case of a biased die. For Laplace, statements of probability must be defined with respect to equally probable cases. But suppose I shift the centre of gravity of this die, or file away one of its corners. If I do this, then it will no longer be the case that each outcome is equally possible. But regardless, we would still like to say that there is some specific probability of throwing an even number with this die. On the classical conception, however,

[6] *Ibid.*

this probability appears to be impossible to calculate.

A more telling example has to do with the probability of death (e.g., as used by insurance companies). Suppose we say that the probability of death for a forty year-old non-smoking male is 0.011. It is not clear at all how the classical interpretation can conceive of a case like this one. Von Mises asks:

> Are there 1000 different probabilities, eleven of which are 'favourable' to the occurrence of death, or are there 3000 possibilities and thirty-three 'favourable' ones? It would be useless to search the textbooks for an answer, for no discussion on how to define equally likely cases in questions of this kind are given.[7]

Probability gets defined, for von Mises, as the relative frequency of an outcome with respect to a sequence of repeatable events. For example, given a sequence of 1000 rolls of a single die, we might observe that 6 turns up 300 times. We can then say that the relative frequency of 6 in this sequence is 300/1000 = 0.3. The longer the sequence, the more closely the observed relative frequency approaches the true probability of rolling a 6 (which is considered to be a 'physical' attribute of the sequence). The probability, then, of getting a 6 with this die, is a kind of idealization: it is defined as the limiting value of the relative frequency of 6 with respect to an *infinite* sequence of rolls.

To illustrate how this works: suppose that after each roll of the die I calculate the relative frequency of 6, rounding off to the first decimal place. I eventually find, after n rolls, that the relative frequency ceases to change; it remains constant at, say, 0.3. At this point, I increase the number of decimal places to two. The value begins to fluctuate again, but after m more rolls, it again ceases to change; it stays constant at, say .32. I then increase the precision to three places. Again, the relative frequency eventually stabilizes; this time at .324. If I continue the process *infinitely*, the relative frequency will be accurate to an infinite number of decimal points.[8]

[7] von Mises 1957, 69-70.

[8] Cf. von Mises 1957, 14-15.

Now on von Mises' view, probability statements must always refer to a 'collective', or what we now normally call a reference class: "a sequence of uniform events or processes which differ by certain observable attributes."[9] One such collective is the sequence of rolls with a biased die, where each member of the sequence potentially differs with respect to the actual number showing on its topmost face at the end of each roll. Another collective is the class consisting of all non-smoking males in their fortieth year of life, where each member of the sequence potentially differs with respect to whether or not they are dead by the end of that year. Specifying an appropriate reference class is of crucial importance in determining relative frequencies. To see why, consider the collective consisting of men over 40, for which we wish to determine the probability of going senile. Relative to this reference class, the probability of senility is likely very low. But if we narrow the reference class so that it consists of men over 90, the probability of senility is much higher. Now for von Mises, probability statements do not have any meaning whatsoever unless they refer to a collective; i.e., on his view there is no such thing as the absolute probability of the event A. All probabilities are conditional probabilities of the form $P(A|R)$, where R is the reference class. So-called 'single-case' probability statements are meaningless, for von Mises.

> It is utter nonsense to say, for instance, that Mr. X, now aged forty, has the probability 0.011 of dying in the course of the next year. ... One might suggest that a correct value of the probability of death for Mr. X may be obtained by restricting the collective to which he belongs as far as possible ... There is, however, no end to this process, and if we go further and further into the selection of members of the collective, we shall be left finally with this individual alone.[10]

A true single case probability, for von Mises, would have to take *this* actual individual as the reference class. And to specify *this* individual we must specify all the ways in which it differs from every other individual. But in that case, there is no sense in speaking of

[9]von Mises 1957, 12.
[10]von Mises 1957, 17-18.

the *probability* of an event with respect to this individual, for we have now defined this reference class so precisely that we can determine the outcome of any future event with certainty. Thus von Mises' frequency interpretation is similar to the classical interpretation at least in this respect: the probability of an event is defined with respect to which properties we choose to include, which properties we choose to ignore, and which properties we have no knowledge of when we define the reference class.

One problem with von Mises' version of frequentism is that it involves an inference from actually observed sequences of events, which are finite, to infinite sequences of those same events. But in fact, as Wittgenstein points out, it is possible to infer infinitely many infinite sequences from a finite sequence (which we take as the infinite sequence's initial segment).

> If we infer from the relative frequency of an event its relative frequency in the future, we can of course only do that from the frequency which has in fact been so far observed. And not from one we have derived from observation by some process or other for calculating probabilities. For the probability we calculate is compatible with *any* frequency whatever that we actually observe, since it leaves the time open.[11]

For example, consider a physical process described by a function that maps to 0 for values of x less than n (where n is some arbitrarily large value), but maps to $2x$ for values of x greater than or equal to n. If we consider only an initial segment of the sequence, we will be led to infer, using the limiting value of the relative frequency as our guide, that the probability of $f(x)$ resulting in 0 is 1. If n is so large that it is impossible for anyone to observe n actual instances of this process, then there is no way to infer the true relative frequency of 0 with respect to the process. Further, infinitely many hypotheses about f are compatible with the relative frequency that we observe. We might hypothesize that $f(x) = 3x; (\forall x \geq n)$, or that $f(x) = 4x; (\forall x \geq n)$, or that $f(x) = 323467x; (\forall x \geq n)$ or even that $(n \leq \forall x < 2n)(f(x) = 3x); (\forall x \geq 2n)(f(x) = 30x)$. All of these hypotheses,

[11] Wittgenstein 1975 §234.

and many others, are compatible with the observed sequence.

This is a serious problem for von Mises, one which Reichenbach avoids by defining probability merely with respect to a *sufficiently long* series of trials:

> If a sequence of roulette results or of mortality statistics were to show a noticeable convergence only after billions of elements, we could not use it for the application of probability concepts, since its domain of convergence would be inaccessible to human experience. However, should one of the sequences converge "reasonably" within the domain accessible to human observation and diverge for all its infinite rest, such divergence would not disturb us; we should find that a *semiconvergent* sequence satisfies sufficiently all the rules of probability. I will introduce the term *practical limit* for sequences that, in dimensions accessible to human observation, converge sufficiently and remain within the interval of convergence.[12]

Reichenbach considers the possibility of redefining the probability calculus with respect to finite sequences; he concludes, however, that while this is possible, "a calculus of this kind would be rather complicated and cumbersome."[13] Thus he regards the definition of the calculus in terms of infinite sequences as an idealization, analogous to the idealization of practical geometry by ideal geometry.

Wittgenstein's Logical Interpretation

Wittgenstein did not write very much on the topic of probability, however the little that he did write[14] has been highly influential in the general development of the so-called *logical* interpretation of probability. The main sources for Wittgenstein's views are propositions 5.1 - 5.156 of the *Tractatus* (his early period), and also §§225 - 237 of the *Philosophical Remarks* (his middle period). Like Laplace, Wittgenstein's notion of probability is epistemic: the probability of a proposition depends on our knowledge situation at the time the

[12]Reichenbach 1949, 347

[13]Reichenbach 1949, 348.

[14]The little that we have of his views on the subject come from a few short pages of the *Tractatus*, some 'remarks' from the 1930's, and informal conversations with the Vienna Circle during the same decade.

statement is made. For Wittgenstein, probability represents a relation between propositions: between the propositions representative of our knowledge and the propositions for which we are seeking to fix a probability value to. It is easiest to understand Wittgenstein's interpretation if we consider a truth table. For example:

	A	B	C	$\mathscr{P}$: $(\sim A \supset (B \vee C))$	$\mathscr{Q}$: $(\sim B)$
1	T	T	T	T	F
2	T	T	F	T	F
3	T	F	T	T	T
4	T	F	F	T	T
5	F	T	T	T	F
6	F	T	F	T	F
7	F	F	T	T	T
8	F	F	F	F	T

Now consider the proposition $\mathscr{P}$. We can view it as a truth function, and like any other function, it accepts a number of arguments as input (here given in the first three columns) and produces a determinate output. Now call the *truth grounds* of a proposition the truth values of its truth arguments which make it true.[15] The rows on which $\mathscr{P}$ is true are the rows 1-7. Therefore the set of its truth grounds is: $\{\ \{T,T,T\},\ \{T,T,F\},\ \{T,F,T\},\ \{T,F,F\},\ \{F,T,T\},\ \{F,T,F\},\ \{F,F,T\}\ \}$. Similarly, the truth grounds of $\mathscr{Q}$ are on lines 3-4,7-8: $\{\ \{T,F,T\},\ \{T,F,F\},\ \{F,F,T\},\ \{F,F,F\}\ \}$.[16] Now if we compare the truth grounds of these two propositions, we find that the truth grounds of $\mathscr{Q}$ and $\mathscr{P}$ overlap. But we can ask the question, 'given $\mathscr{P}$, when is $\mathscr{Q}$ true?' $\mathscr{P}$ is true on lines 1-7 for a total of seven instances. In the context of these seven lines, $\mathscr{Q}$ is true three times (on lines 3, 4, and 7). So we can say that $\mathscr{Q}$ is true 3/7 of the time whenever $\mathscr{P}$ is true; i.e., we can calculate the conditional probability: $P(\mathscr{Q}|\mathscr{P}) = 3/7 \approx .4286$.

In general, we can say that: "If T_r is the number of the truth-grounds of a proposition 'r', and if T_{rs} is the number of the truth-grounds of a proposition 's' that are at the same time truth-grounds of 'r', then we call the ratio $T_{rs} : T_r$ the degree of *probability* that

[15]Wittgenstein 2005, 5.101.

[16]Strictly speaking, $\mathscr{Q}$, takes only one argument. Since the truth-values of A and C are irrelevant, however, including them has no effect.

the proposition 'r' gives to the proposition 's'".[17] When two propositions have no truth arguments in common, the probability of one given the other is simply 1/2. For example, A, B, and C above have no truth arguments in common, and we can see from the truth table that $P(A|B) = P(B|A) = P(A|C) = P(C|A) = P(B|C) = P(C|B) = 0.5$.

The truth-grounds of a proposition can be said to define a proposition's range. Wittgenstein writes, "The truth-conditions of a proposition determine the range that it leaves open to the facts. ... "[18]. Consider the truth table above once again, and imagine that the totality of our knowledge consists of the elementary propositions A, B, and C. We can say, then, that each line of the truth table represents a possible state description of the universe. For example, line 6 represents a possible state of the universe such that $(\sim A\&B\& \sim C)$. We can now define the *range* of the proposition $\mathscr{Q}$ as the state descriptions that are compatible with it. Further, we can say that $\mathscr{Q}$ is logically equivalent to the disjunction of the state descriptions making up its range, i.e.: $(A\& \sim B\&C) \vee (A\& \sim B\& \sim C) \vee (\sim A\& \sim B\&C) \vee (\sim A\& \sim B\& \sim C)$

Now for the Wittgenstein of the *Tractatus*, all non-elementary propositions are logically analyzable in principle into truth-functions of elementary propositions. Thus if we could perform such a complete analysis, then we could compute the probability of some proposition of unknown truth value simply by comparing its range with the total number of state descriptions that are left open by the other propositions which make up our knowledge situation. Thus imagine, for instance, that we know independently that the state description described by line 1 can be ruled out. Then in the case where we do not know the truth value of $\mathscr{P}$, we can say that its probability given our knowledge situation is 6/7: it is true six out of seven times (i.e., true on lines 2-7 but false on line 8). Note how Wittgenstein's conception is similar to Laplace's. Both view probability as a ratio of 'favourable' cases to total possible cases. For Laplace, these are possible outcomes; for Wittgenstein, these are truth grounds. But unlike Laplace's conception, which is absolute, Wittgenstein's ba-

[17]Wittgenstein 2005, 5.15
[18]Wittgenstein 2005, 4.463

sic conception of probability is *conditional.* Wittgenstein has no need of the principle of insufficient reason. Whether a statement of unknown truth value is true or false is not simply assumed to be equally likely - its possibility is determined by its relation to the other statements making up our knowledge situation. Thus for Laplace, we 'divide our ignorance' into equally possible cases expressing absolute probabilities. For Wittgenstein, this possibility itself is conditional upon our *knowledge.*

One might question the usefulness, however, of an interpretation of probability that requires us to completely analyze our propositions into elementary propositions. While this may be acceptable for an idealized logical language, the analysis of propositions in natural or even scientific language is notoriously difficult.[19]. However, we can address this concern if we consider the following example. Wittgenstein asks us to imagine the case of an urn into which we place an equal number of white and black balls. If I now ask, "what is the probability of drawing a white ball from this urn?", the answer will invariably be 1/2. Why? Wittgenstein's answer is the following "... if I say, 'The probability of my drawing a white ball is equal to the probability of my drawing a black one', this means that all the circumstances that I know of (including the laws of nature assumed as hypotheses) give no *more* probability to the occurrence of the one event than to that of the other. ...".[20]

The key phrase here, ironically, is the one in parentheses.[21] Thus imagine the set of all (not necessarily elementary) propositions whose truth value is known. Call this set the bulk of our knowledge, $\mathscr{K}$. Now, for Wittgenstein, the proposition $(P \supset (Q \vee R))$ can belong to $\mathscr{K}$ even though neither P nor Q nor R belong to $\mathscr{K}$. This may seem strange at first, but that this is possible follows from Wittgenstein's views on quantification. For Wittgenstein, quantified sentences are not propositions; rather, they are merely schemas or prototypes for constructing propositions. For example, he writes: "I dissociate the concept *all* from truth-

[19]Cf. Black 1966, 256.

[20]Wittgenstein 2005, 5.154

[21]My interpretation is similar to that given in von Wright, 1969. However, I disagree with von Wright's negative evaluation of Wittgenstein's views on the probability of hypotheses.

functions ...";[22] and also: "What is peculiar to the generality-sign is first, that it indicates a logical prototype, and secondly, that it gives prominence to constants";[23] "The generality-sign occurs as an argument.";[24] "An hypothesis is a law for forming propositions."[25]

For example, the proposition: 'If Socrates is a man, then Socrates is mortal' is an instance of the rule specified by the hypothesis: 'All men are mortal'. But for Wittgenstein, only the former sentence is truly a proposition. The latter is merely a schema. Thus if we hold a certain law of nature to be true, e.g., $(\forall_x)(A_x \supset B_x)$, then this law itself will not belong to $\mathscr{K}$; however, all propositions capable of being constructed from it, e.g.: $(A_a \supset B_a)$, $(A_b \supset B_b)$, $(A_c \supset B_c)$, ..., will belong to $\mathscr{K}$; and this will be the case even if the truth-values of $A_a, A_b, A_c, B_a, B_b, B_c, \ldots$ are unknown.

Thus it is possible to have knowledge of a non-elementary proposition without having knowledge of its constituents; for the knowledge of such propositions is *derived* from our general laws and not *built up* from elementary propositions. The hypothetical general laws provide a means to add propositions expressing relations between propositions (elementary or not) to $\mathscr{K}$ without actually adding knowledge of these propositions themselves to $\mathscr{K}$. But if we can 'know' a proposition without knowing the truth or falsity of its constituents, then we can compute probabilities based on these unanalysed propositions as well. This is Wittgenstein's meaning when he writes: "It is in this way that probability is a generalization. It involves a general description of a propositional form. We use probability only in default of certainty - if our knowledge of a fact is not indeed complete, but we do know something about its form. ..."[26]

Coming back to the urn example, say that it is a law of nature that 'for all urns containing an equal number of white and black balls, all balls drawn will be either white or a black' (where 'or' is, of course, taken in the exclusive sense), or symbolically: $(\forall_{x \in U})(W_x \oplus B_x)$.

[22] Wittgenstein 2005, 5.521.
[23] Wittgenstein 2005, 5.522.
[24] Wittgenstein 2005, 5.523.
[25] Wittgenstein 1975, §228.
[26] Wittgenstein 2005, 5.156.

This implies that for this particular draw, that $(W_a \oplus B_a)$. Now consider the proposition expressing the fact that I draw a white ball, (W_a). The probability conferred on this proposition by $(W_a \oplus B_a)$ is 1/2 (all else being equal). Note that we have calculated this probability in spite of the fact that we do not know either W_a or B_a.

Note also that it is not necessary for either W_a or B_a themselves to be elementary propositions. These, in turn, may be complex, representing a conjunction, perhaps, of elementary propositions. The laws of nature (which are simply assumed, a priori), tell us something about the relationships between propositions, and in describing these relationships, they of course must describe the constituents of these relationships - but these need not be described in every detail for we do not *need* to know what the constituents of these constituents are; these constituents can represent 'abstractions', in a sense. Thus, these laws of nature express *relationships* between (possibly complex) propositions. These laws constitute the framework by which we describe the world;[27] they form the 'a priori' assumptions that we hold in the background when we make any statement of probability. Of course, as our knowledge grows, we may learn more about the constituents of these propositions, and this will, in turn, affect our calculated probabilities. Thus assume that $\mathscr{K}$ consists of the single proposition, $(P \supset (Q \vee R))$. Given that I know this proposition, but not P, Q, or R, the probability that this proposition confers on the proposition R is 4/7. Now assume that as our knowledge grows we eventually add Q to $\mathscr{K}$. In this new situation, the probability of R becomes 1/2. And if, one day, we discover that Q is capable of further analysis, and we come to know the truth or falsity of one of its constituents, say, q_1, this will affect the probability of R even further.

Now as we have seen in the previous section, observed relative frequencies can have no direct bearing, for Wittgenstein, on the calculation of probabilities. By themselves, they show nothing, for as he says, they 'leave the time open'. Relative frequencies can have a role to play with regard to probabilities, however, in an indirect way; i.e., they can

[27]Cf. Wittgenstein 2005, 6.342

inspire us to formulate new hypothetical laws of nature.[28] But aside from a warning that our hypotheses should not be *ad hoc*[29] (i.e., not directly copied from statements of relative frequency) Wittgenstein says nothing with regards to how this 'inspiration' is supposed to work. For this, we turn to Carnap.

Carnap's Probability$_1$ and Probability$_2$

Carnap's views on probability are best described as an attempt at reconciling the frequency and logical interpretations. He observes that:

> It has repeatedly occurred in the history of science that a vehement but futile controversy arose between the proponents of two or more explicata who shared the erroneous belief that they had the same explicandum; when finally it became clear that they meant different explicanda, unfortunately designated by the same term, and that the different explicata were hence compatible and moreover were found to be equally fruitful scientific concepts, the controversy evaporated into nothing.[30]

For Carnap, the two conceptions of probability are defined as: (i) probability$_1$: the degree of confirmation of an hypothesis. (ii) probability$_2$: the relative frequency in the long run. Both probability$_1$ and probability$_2$ are thought of as functions; each of which take a distinct type of argument: probability$_1$ is a function operating on *sentences* (which express propositions), while probability$_2$ is a function that operates on "properties, kinds, classes, usually of events or things."[31] Since probability$_1$ operates on sentences in a language, it does not, strictly speaking, say anything about the actual facts. Elementary statements of probability$_1$ are determined to be true or false solely through logical analysis (this is analogous to Wittgenstein's conception of probability as a relation between propositions). Nevertheless, it is important for probability$_1$ statements to *refer* to the facts, via sentences

[28]Cf. Wittgenstein 1975, §229.
[29]Cf. McGuinness 1979, 95.
[30]Carnap 1962, 26.
[31]Carnap 1962, 33.

describing a body of evidence. A probability$_1$ statement is, in fact, a measure of the support that the evidence, e, gives to the hypothesis, h. That is, when we say that the probability$_1$ of h with respect to e is high, we mean that e gives strong support to h. Symbolically, this is expressed as $c(h, e)$, or in words: the degree of confirmation of h, a hypothesis, with respect to the evidence e.

Now we can quantify a probability$_1$ statement, according to Carnap, in terms of a 'fair betting quotient', and this is how the link between probability$_1$ and probability$_2$ is drawn.[32] Consider the following, somewhat light-hearted example: suppose there are two persons, call them Eric and Linda, and suppose that they agree to bet on whether a cat (call him Silvy) is or is not a six-toed cat. Now suppose further that Eric and Linda have unlimited funds and that Eric bets u_1 and Linda bets u_2 (where $u_1, u_2 > 0$). Call $q = u_1/(u_1 + u_2)$ the *betting quotient*. If we let e be the information concerning Silvy that both Eric and Linda share (that in spite of his name, he is a tom-cat), then we say that probability$_1 = q$ if Eric and Linda's bet is a *fair* bet, given e.

Now to clarify the notion of a fair bet, imagine that Eric and Linda make a whole class of similar bets on each member of the set of tom-cats with female-sounding names (of which Silvy is merely one). Now suppose, for the sake of argument, that there are 100 such cats, and that the *actual* relative frequency, r, of six-toed cats in this population is .25. In this case, Eric will win $rn = 0.25 \times 100 = 25$ of these bets and lose the remaining $(1 - r)n \times 100 = 75$ of them. Say that Eric wagers \$1.00 on each cat, and that Linda wagers \$3.00. Eric will win a total of $rnu_2 = 25 \times \$3 = \$75.00$ from Linda. He will lose the remaining 75 bets, of course, for a loss of \$75.00 (he bets \$1.00 on each). Thus after all the betting is done, both Eric and Linda will break even, and this represents the situation in which a bet on any particular cat in this population is a fair bet. To see why we call this a *fair* bet, suppose that we *tell* Eric and Linda the actual value of r (i.e., 0.25). Would it be rational, now, for Eric to wager \$2.00 on each cat? In this case, $q = \frac{2}{2+3} = 0.4 > r$. He

[32]Carnap 1962, 165-168.

will still win $rnu_2 = \$75.00$ from Linda, but he will now lose $75 \times \$2 = \$150.00$ for a net loss of \$75.00. In general, when $q > r$, Eric will refuse to make such a bet, for such a bet is not fair. Similarly for Linda, she will refuse to make a bet if $q < r$. It is only when $q = r$ that we call this a fair bet. But now recall that probability$_1 = q$ when the bet is fair, and that probability$_2 = r$. Thus in the case of a fair bet, probability$_1$ = probability$_2$.

Now suppose that neither Eric nor Linda knows r (i.e., that r is not contained in e).[33] In order to arrive at a fair betting quotient, Eric and Linda will require that the probability$_1$ of Silvy being six-toed represents an *estimate*, r', of the actual unknown relative frequency, r, of six-toed cats, given the evidence, e. And just as for the case when the actual relative frequency is known, neither Eric nor Linda will make a bet unless $q = r'$, where r' is the best estimate available of r given e. I will end here and will not go into the details of Carnap's procedure for the actual estimation of the relative frequency given the available evidence. This would take us deep into the problem of induction and that is not the subject of my thesis. But the foregoing should be sufficient for elucidating the relationship which, according to Carnap, holds between the two conceptions of probability.

> The relation between probability$_2$ and probability$_1$ is hence seen to be a special instance of the logical relation which holds generally between an empirical, e.g., physical, quantitative concept and the corresponding inductive-logical concept of its estimate with respect to given evidence. This relation explains, on the one hand, the different nature of the two probability concepts, but, on the other hand, also the far-reaching analogy between them which we shall repeatedly observe in our further discussions.[34]

Popper's Propensity Interpretation

Popper originally ascribed to a frequentist interpretation of probability, but eventually came to reject this view in favour of the propensity interpretation, and he gives us the following consideration in order to show why. Consider our old friend again, the biased

[33]Cf. Carnap 1962, 168-175.
[34]Carnap 1962, 173.

die, and imagine that, after many experiments, we come to determine the relative frequency of 6 for the sequence of rolls with this die to be 1/4. Now imagine that we modify the sequence slightly: suppose that it consists of a large number of throws with the biased die, mixed in with a few throws (two or three) with an unbiased die. Now the probability of rolling a 6 for each of these few throws with the unbiased die is clearly 1/6, even though "these throws are, according to our assumptions, *members of a sequence* of throws with the statistical frequency 1/4.[35]

Now the obvious response is that when we say that the probability of 6 with respect to the unbiased die is 1/6, we have simply narrowed the reference class. It is analogous to saying that the probability of senility for men over 40 is 0.000003, while it is 0.03 for men over 90 (I am making up these numbers, of course). With respect to the original sequence we simply choose not to take account of the physical differences of the two dice when defining our reference class. But we know, say, based on the observation of previous experiments with an unbiased die, that the probability with respect to this sequence is 1/6. Thus we can say that $(i)\quad p(\text{‘6’}|b) = 1/4$ (where b is the mixed class), but that $(ii)\quad p(\text{‘6’}|bc) = 1/6$ (where c is the subclass of throws with a correct die). Now Popper writes:

> Of course there is no doubt as to the compatibility of the two equations ... nor is there any question that these two cases can be realised within the frequency theory: we *might* construct some sequence b such that equation (i) is satisfied, while in a selection sequence bc - a very long and virtually infinite sequence whose elements belong both to b and c - equation (ii) is satisfied. *But our case is not of this kind.* For bc is not, in our case, a virtually infinite sequence. It contains, according to our assumptions, at most three elements. In bc the six may come up not at all, or once, or twice, or three times. But it *certainly* will not occur with the frequency 1/6 in the sequence bc because we *know* that this sequence contains at most three elements.[36]

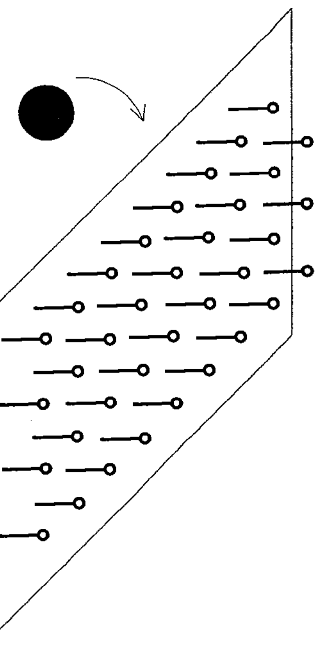
Thus in order to determine that $p(\text{‘6’}|bc) = 1/6$ we must interpret c as being a long

[35]Popper 1959. 31-32
[36]Popper 1959, 33.

sequence of throws with a correct die, and "we estimate or conjecture that, in a sequence of throws with a correct die, the six will come up in 1/6 of the cases."[37] According to Popper, this forces the frequency theorist to introduce a slight modification to his interpretation stipulating what can count as an admissible sequence. The frequency theorist will have to say that only those sequences "which are *characterized by a set of generating conditions* - by a set of conditions whose repeated realisation produces the elements of the sequence"[38] are admissible. In this case, *b* is no longer an admissible sequence. "Its main part, which consists only of throws with the loaded die, will make an admissible sequence, and no question arises with respect to it. The other part, *bc*, consists of throws with a regular die, and belongs to a virtual sequence *c* - also an admissible one - of such throws. There is again no problem here. It is clear that, once the modification has been adopted, the frequency interpretation is no longer in any difficulty."[39] But if we make this modification, then we have made the transition from the frequency interpretation to the propensity interpretation; for if the sequence is defined by the set of its generating conditions, probability "may now be said to be *a property of the generating conditions*."[40] And if we say that, then we must view the generating conditions as having a certain disposition, or propensity to generate sequences with frequencies equal to the probabilities.

For Popper, a probability statement is interpreted as a real physical tendency, or propensity - we could also call it a potential - for the experimental setup under investigation to give rise to a particular result. Given that a propensity is a real property of the experimental conditions (analogous to a force), probability statements express propensities that can influence experimental results, interfere with one another, and most importantly: they can be objectively tested. In Popper's colourful language, a propensity can be 'kicked', and it can 'kick back'. He illustrates this by means of a 'pinboard',[41] i.e.:

[37] Popper 1959, 33.
[38] Popper 1959, 34.
[39] Popper 1959, 34.
[40] Popper 1959, 34.
[41] Popper 1982, 72.



Say, for example, that we throw a number of balls down this (symmetrical) board. After a sufficient number of trials, the balls will accumulate at the bottom of the board and form a normal distribution curve, which we will say represents the probability distribution curve for a single experiment. Now suppose we 'kick' the board by slightly tilting it counter-clockwise. In this case, we will also kick the propensities: the probability distribution for a single case will be biased in favour of the left end of the board; and the propensities will 'kick back': the shape of the pile at the bottom of the board will be different. On the other hand, suppose we remove one of the pins. This will again alter the probability distribution curve - and since the curve represents the probability distribution for each ball that rolls down, it will alter the probability *for every single experiment*, whether or not the ball comes close to the removed pin. We can also ask, for example, what the probability distribution curve is for a ball, given that it hits a specific pin on its way down. Importantly, any hypothesis we make about the distribution curve can be *tested* by means of repeated experiments. I will have more to say about the propensity interpretation below.

2.2 Hempel's Inductive-Statistical Model of Explanation

But let us come back, now, to our primary topic: explanation, and let us now consider the case of probabilistic or statistical explanation: where we wish to explain the occurrence of a phenomenon that we cannot (although perhaps in principle we could) predict with certainty. Now just like the case of D-N explanation, Hempel's Inductive-Statistical (I-S) model takes the form of an argument; however the inference in this case is *inductive*, not deductive. We can represent I-S explanations as follows:

$$\begin{array}{l} p(R,S) \text{ is close to } 1 \\ \underline{S_j} \\ \hline\hline R_j \end{array} \quad \text{[makes practically certain]}$$

Here, $p(R,S)$ is a sentence expressing a probabilistic law. The probability expressed in this sentence is a *statistical* probability, interpreted according to the frequentist interpretation. It asserts that the relative frequency of the property R with respect to the population S tends toward 1 in the long run. Unlike the case for D-N explanations, the explanans statements in I-S explanations do not logically entail the explanandum statement. Rather, they confer a *degree of confirmation* on the explanandum statement (represented by the double-line). The degree of confirmation expresses a *logical* probability.[42] Logical probability, as we saw above, is a relation between sentences. It is the degree of confirmation, or inductive support, conferred on a hypothesis, H, by the body of knowledge, K, and it is represented by $c(H,K)$. In the case of I-S explanation, this relation expresses the extent to which the explanandum statement is confirmed by the explanans. Thus, another way of representing the I-S schema would be as follows:

$$\begin{array}{l} p(R,S) = r \\ S_j \\ \hline c(R_j, S_j) = r \end{array}$$

[42] Cf. Hempel 1966, 310.

Strictly speaking, an I-S argument is not an argument for the explanandum statement *itself*, but merely an argument for the degree of inductive support that we have for it. We are not permitted to detach E. We take the extra step and say that an I-S argument is a good *explanation* for the explanandum statement when the degree of confirmation is high (≈ 1); i.e., when the explanandum is to be expected given the explanans.

For example, let S_j represent the fact that Joanne Jones has a streptococcus infection, P_j the fact that Joanne has been administered penicillin, and R_j the fact that Joanne has recovered from her illness. If the relative frequency of R with respect to S and P tends toward the limit of, say, .95 in the long run, then we can give an I-S explanation of Joanne's actual recovery as follows:

$$\begin{array}{l} p(R, S \cdot P) = .95 \\ \underline{S_j \cdot P_j \qquad\qquad\qquad\qquad} \\ \hline R_j \end{array} \; {}_{.95}$$

But consider a variation on the foregoing example: suppose that Joanne has a penicillin-resistant strain of streptococcus, S^*; thus, her probability of recovering is actually close to 0, and we can present an I-S argument for her non-recovery ($\overline{R}$) like so:

$$\begin{array}{l} p(\overline{R}, S^* \cdot P) \text{ is close to } 1 \\ \underline{S_j^* \cdot P_j \qquad\qquad\qquad\qquad} \\ \hline \overline{R}_j \end{array} \; \text{[makes practically certain]}$$

We now have two arguments, both of which have true premises (which are mutually consistent), that confer high inductive support to mutually contradictory conclusions. Whether Joanne recovers or not, we can argue that the outcome was to be expected. But such a result serves to make all so-called I-S 'explanations' seem rather dubious, and it makes them utterly useless for the purposes of prediction.

Hempel's solution is the requirement of maximal specificity (RMS). According to RMS, the probabilistic laws invoked must be specified relative to the most specific reference class available given a knowledge situation, K. Thus, an I-S argument for Joanne's recovery that does not take into account the fact that her infection is resistant to penicillin is not sufficient to explain her recovery.

More formally: let s be the conjunction of the statements in the explanans. If there is a sentence, k, which is entailed by our knowledge situation, K, such that $s \cdot k$ implies that the case under consideration belongs to a subset, F_1, of the original reference class, F, and if $p(G, F_1) \neq p(G, F)$, then RMS requires us to consider the more restricted reference class, F_1, unless $p(G, F_1)$ amounts to a theorem of probability theory (the latter qualification is there to guard against assigning Joanne to the reference class of persons who have recovered from streptococcus infections in order to explain the fact that she has recovered from her streptococcus infection. In that case, we would have $p(R, R \cdot P)$, which is, of course, trivially equal to 1 (recall, also, von Mises' criticisms of single-case probabilities)).[43]

But there are additional problems. Consider the following case.[44] Say that a teaspoon of salt has a .95 probability of completely dissolving after five minutes when stirred into a bowl of cold water. Now imagine that I do not know this, and that I form the hypothesis that all salt which has been placed under a magician's spell has the probability .95 of dissolving in a similar manner. I perform experiments on various samples (using various magicians). Lo and behold, my hypothesis is confirmed: the relative frequency of dissolved samples tends toward the limit of .95 in the long run. I then check for a more specific reference class. I find none that affect the relative frequency. I therefore feel confident that the salt sample dissolved because it was hexed. Surely, however, the statistical law is irrelevant to the explanation in spite of its being confirmed by the evidence, and in spite of there being no more maximally specific reference class.

Another problem is I-S's inability to explain low-probability outcomes. For example, in quantum mechanics, there is a known law for the probability of a single instance of alpha decay in a U^{238} atom. It seems as though we can explain why alpha decay occurs by

[43]See: Hempel 1965, 400.
[44]Cf. Curd & Cover, 786.

appealing to this law. But on Hempel's account, since the probability of alpha decay is so small, the degree of confirmation we can confer on the explanandum is not high enough to be used in an I-S explanation of it. Thus alpha-decay cannot cannot be explained at all.

2.3 Probabilistic Explanation According to the D-N-P Model

Let us consider how the D-N-P model handles these examples. With regards to the irrelevance cases, Railton makes it a requirement that the probability laws must be true; i.e., that they describe genuinely indeterministic processes (statistical generalizations are ruled out) that give rise to the phenomenon, and that they are derivable from our best scientific theories.[45] Thus since the mechanism is responsible *for* the phenomenon, it is obviously relevant *to* the phenomenon. The problem of maximal specificity also does not arise, for the laws, by their very nature, are maximally specific (they *truly* describe the phenomenon in *every* respect).

As for the low probability cases, recall that for Railton, the requirement is not to give a good reason to expect, but to give an account of the mechanism *which gives rise* to an outcome; an explanation of a phenomenon is an explanation of *how it came about.* Recall the schema of a D-N-P explanation:

(1) A *complete* theoretical account of the mechanism involved in the production of the phenomenon described by the explanandum statement, and a derivation, from our theory, of the laws required to infer the probability of occurrence of the explanandum phenomenon, E.

(2) A D-N argument from the laws derived in (1) to the probability of occurrence of E.

(3) A statement to the effect that E actually did occur.[46]

The role of the D-N portion of the D-N-P text is merely to infer a probability for the event to occur. But unlike the I-S model, the probability inferred cannot be understood as

[45]Cf. Railton 1978, 222.
[46]Cf. Railton 1978, 214, 218.

the degree of confirmation. Its role is not to show that the explanandum event was to be expected; rather, its role is to elucidate the probabilistic *mechanism* behind the occurrence of the phenomenon. Thus a probabilistic D-N-P explanation shows us how it is that there is a particular probability for a phenomenon to occur, and then by means of the parenthetic addendum, tells us that, by the way, it did occur. But now in order to show how the mechanism *gave rise* (however improbably) to the the explanandum event, we need to interpret the probability that we infer as a physical probability - in other words, as a *propensity*. A propensity, as we saw above, is a *real* property of a physical system. It is the disposition of one particular state of a physical system to give rise to a later state of the same system. Thus if I say that the probability that a die turns up 6 on a particular roll is 1/6, I am saying that the propensity for the die to turn up 6 is 1/6; I am saying that it is a real property of this die in this instance[47] that it has an equal chance of landing on any face. Railton needs to interpret single-case probabilities as propensities because rather than simply saying that the low-probability occurrence of the explanandum event is inexplicable, he wants to say that the explanandum event is somehow tied to the inferred probability - that it results from it - and a propensity interpretation of probability can give us this link. For on the propensity interpretation, these probabilities can *cause* the explanandum event:

> Under a *propensity interpretation*, probability has the characteristics sought: a probability is the expression of the strength of a physical tendency in an individual chance system to produce a particular outcome; it is therefore straightforwardly applicable to single cases; and it is (in a relevant sense) causally responsible for that outcome whenever it is realized.[48]

Propensities are certainly able to do the job that Railton needs them to do. However, there are certain problems. For one, a propensity interpretation of probability is probably incompatible with the Copenhagen interpretation of quantum mechanics. Indeed, part of the motivation for the propensity interpretation in the first place was what Popper saw as the

[47] Strictly speaking it is a property of the physical system, including my hand, the table, etc., of which the die is merely a part.

[48] Railton 1978, 222.

problems with the Copenhagen interpretation: "I had always been convinced that the problem of the interpretation of the quantum theory was closely linked with the problem of the interpretation of probability theory in general, and that the Bohr-Heisenberg interpretation was the result of a subjectivist interpretation of probability".[49] As I will explain in much more detail in the next chapter, the irreducibly probabilistic nature of atomic phenomena results, on the Copenhagen interpretation, from the impossibility of fully separating the measuring instrument from the object being measured in our characterization of the results of experiments. On the Copenhagen interpretation, this means that the role of the observer - a subjective element - can never be fully eliminated in the atomic domain (unlike the case for the macro domain of classical physics). But on Popper's view the uncertainty is interpreted, not as resulting from the intrusion of the subject, but as a real property of the experimental setup. "Probability fields are physical, even though they depend on, or are relative to, specified experimental conditions."[50] Further, the wave-particle duality of atomic phenomena which, as we will see in the next chapter, is at the heart of the Copenhagen interpretation is rejected by Popper. On Popper's view, a particle description of atomic phenomena is actually the only one that is valid. The 'waves' are seen as probability fields, interpreted as propensity fields.

> Of course they may in their turn become objects of study, just like other physical properties, such as the momentum of a particle. As we have seen, they may even be objects which one can kick and which can kick back. There is, however, no duality in any sense in which we may speak either of a particle or of a wave but not both at once - exactly as there is no duality between a particle and its momentum that prevents us from speaking of both at the same time.[51]

I am not going to undertake a defence the Copenhagen interpretation of quantum mechanics here, but I will note that Railton, defending the view that it is possible to explain irreducibly probabilistic processes, writes:

---

[49]Popper 1959, 27-28

[50]Popper 1982, 81.

[51]*Ibid.*

> I take it that quantum mechanics, under the dominant interpretation, gives us reason to think that there are irreducibly probabilistic processes in nature, and that they are governed by probabilistic and non-probabilistic laws. ... If the dominant interpretation of quantum mechanics is right, there are no "hidden variables" characterizing unknown initial conditions of nuclei that suffice, in conjunction with deterministic laws, to account for the occurrences of alpha-decay ...[52]

Unless Railton is mistaken with regard to what the dominant interpretation of quantum mechanics is, I take it that he means the Copenhagen interpretation. If so, then given that almost all of the examples that Railton uses to illustrate his D-N-P model have to do with atomic processes, we must question whether he is actually being consistent or not. One might answer, of course, that Popper's propensity interpretation of *quantum mechanics* aside, perhaps the Copenhagen interpretation is actually compatible with the propensity interpretation of *probability*. However the propensity interpretation of probability itself, as Railton admits[53], is notoriously difficult. It is not clear, exactly, what the relationship between propensities and the probability calculus is. As Humphreys has shown,[54] propensities are asymmetrical while the probabilities used in the traditional probability calculus are not. Since propensities are causal; it makes no sense to speak of the propensity for a state *to have arisen* from a previous state. Humphreys suggests, in fact, that we throw out the probability calculus and start anew with propensities; I think, however, that this is a touch extreme. Further, a marriage of the Copenhagen interpretation of quantum mechanics with the propensity interpretation of probability would be most awkward, to say the least, for as I will explain in the next chapter, on the Copenhagen interpretation of quantum mechanics, it makes no sense at all, strictly speaking, to describe atomic phenomena causally.

I do not think it is necessary to commit ourselves to a propensity interpretation of probability for the purposes of giving an account of scientific explanation. Propensities allow us to explain why a particular atom improbably undergoes alpha-decay (it is 'because' it

[52]Railton 1981, 234.
[53]Railton 1978, 222.
[54]Cf. Humphreys 1985.

had a propensity, however small, to do so). But what do we really lose if we give this up? We have, after all, explained why there was such-and-such a probability for this event to occur, and using this fact we *can* explain why *this particular sample* of Uranium has such-and-such a half life. Why not just say that we are unable to explain, causally at least, *this particular* occurrence of alpha-decay? As Kitcher writes (referring to the Schrödinger wave equation): "Conjoin as many as you like of the propositions that occur in this derivation or consider the entire derivation. Whatever your choice, you will not have shown why e_2 tunnelled through, rather than being reflected, or why e_2, rather than e_1, tunnelled through. For what those why-questions ask is for a specification of the differences between electrons that tunnel through and electrons that are reflected, and it is, of course, part of character of [quantum mechanics] that there are no such differences to be found".[55] At the very least we should hedge our bets. *If* propensities can be made clear, and *if* they can be applied to quantum mechanical phenomena, so much the better for our account of explanation. Until such time, however, I think it is best to suspend judgement. This does not undermine, to my mind, the D-N-P. For again, we have explained why the atomic event had the probability to occur that it did.

We are coming near, now, to the end of our discussion. It will be useful again, before moving on, to survey the ground we have covered so far. I argued in the first chapter that Railton's D-N-P model of explanation holds great promise as a model capable of accommodating both causal and pragmatic concerns, but that its commitment to hard reductionism was unnecessary and could be dropped. In this chapter, I showed how Railton's D-N-P model is able to overcome many of the counter-examples to Hempel's theory, and I explained that due to his desire to account for explanations of improbable events, Railton is forced to embrace a propensity interpretation of probability. I argued, however, that an account of explanation should not commit itself to such a controversial thesis - especially one that seems to be at odds with much of the explaining that quantum physicists actually do.

[55]Kitcher 1989, 451.

The price of dropping the propensity interpretation may be that we have to forego causal explanations of irreducibly indeterministic phenomena, but I argued that this price is not so steep as it may seem. Further, as I will show in detail in the next chapter, there are no causal explanations of irreducibly indeterministic phenomena, at least not according to the dominant interpretation of quantum mechanics.

But if the Copenhagen interpretation is incompatible with a causal interpretation of atomic phenomena, then does this not undermine the basic motivation behind the D-N-P model? Recall that the principal requirement of the D-N-P model is to elucidate the *mechanism* behind the occurrence of some phenomenon, and when one speaks of a mechanism one usually speaks of a *causal* mechanism. But as I pointed out above when I introduced the D-N-P model, Railton himself does not actually impose a causal requirement.

> ... it will not do simply to add to the D-N model a requirement that the explanans contain causes whenever the explanandum is a particular fact. First, some particular facts may be explained non-causally, e.g., by subsumption under structural laws such as the Pauli exclusion principle. Second, even where causal explanation is called for, the existence of general, causal laws that cover the explanandum has not always been sufficient for explanation: the search for explanation has also taken the form of a search for mechanisms that underlie these laws.[56]

My disagreement with Railton is not over whether or not there are non-causal explanations. Both of us agree that there are. Rather, my disagreement with Railton is over the *extent* to which non-causal explanations exist in science. Railton holds that explanations of irreducibly probabilistic processes (i.e., those given in quantum mechanics) are actually causal in nature; thus he requires a propensity interpretation of probability in order to to interpret those low probabilities *objectively* and to thus have a notion of the cause of an improbable atomic event. I, on the other hand, deny that explanations of atomic phenomena are causal (at least according to the Copenhagen interpretation), and therefore I deny that we need to commit ourselves to a propensity interpretation of probability. But whatever

[56]Railton 1978, 207.

view one takes on the causal or non-causal status of atomic phenomena, the D-N-P (minus the commitment to propensities) is compatible with that view.

Chapter 3

Explanations of Atomic Phenomena

This last chapter will take the form of an extended example. I will describe the characterization of atomic processes as given by the dominant interpretation of the most successful theory in physics: quantum mechanics. I will show how, according to the Copenhagen interpretation of quantum mechanics, causality does not strictly speaking apply to atomic phenomena. As well, I will show how the Copenhagen interpretation presents us with one more reason to reject the extreme reductionistic view of science that I argued against in chapter 1.

3.1 The Genesis of Quantum Mechanics

When we subject certain objects to extremely high temperatures, they absorb energy and emit light (picture a blacksmith hammering metal in a forge; if the temperature is high enough, the metal becomes 'red hot' or 'white hot'). In order to study this phenomenon, 19^{th} century physicists conceived of the notion of a 'black body' - a sort of idealization: a black body is a completely non-reflecting object that is capable of emitting radiation at an intensity directly proportional to the amount of energy it absorbs. Finding a theoretical model for this so-called 'black body radiation' was the problem that was to lead physicists down the road toward what is now known as quantum mechanics. At around the middle of the century (1859-60), Kirchoff showed that the ratio of emitted to absorbed energy in these materials depended solely on the frequency and temperature, and not, for example, on the shape of the object or on the nature of the material from which it is made. He left it as a challenge for subsequent physicists to find a precise expression for the relationship between these quantities.[1]

At the turn of the century, Planck narrowed down the problem by focusing attention not on the radiation *per se*, but on the radiating atoms making up the material (the so-called 'oscillators'). He was able to devise a formula for the frequency-energy relationship, but it was one that was based on some very ad hoc assumptions;[2] while the formula successfully described the experimental results, there was still the problem of finding an underlying physical basis for it. Planck approached this problem by borrowing ideas from Boltzmann's work in statistical thermodynamics: he hypothesized that the total energy of the system was distributed over a large collection of indistinguishable energy elements (oscillators). The result was the now famous formula: $\varepsilon = h\nu$ - i.e., the energy, ε, in each element is equal to the constant of proportionality, h (Planck's constant), times the frequency, ν. Curiously, his results implied that ε must be given in *integer multiples* of $h\nu$. This result was unprecedented. Heisenberg speculates:

> ... he must soon have found that his formula looked as if the oscillator could only contain discrete quanta of energy - a result that was so different from anything known in classical physics that he certainly must have refused to believe it in the beginning. ... Planck must have realized at this time that his formula had touched the foundations of our description of nature ...[3]

The discovery of Planck's constant (or the 'quantum of action' as it came to be known) was key to the resolution of some of the other outstanding problems in physics at the time. Thus as a result of his work on the photoelectric effect, Einstein, in 1905, was able to describe light in terms of energy quanta, and in 1907, as a result of his work on the specific

[1]Cf. Baggott 2004, 10.
[2]Cf. Baggott 2004, 12.
[3]Heisenberg 1959, 35.

heats of solid bodies, he did the same for atoms and ions. In 1913 Bohr published his theory of the atom, according to which electrons are confined to fixed orbits around the nucleus of an atom that depend on integer numbers counted outwards from the nucleus (the so-called 'principal quantum numbers').

By 1924, de Broglie was able to relate Planck's relation $\varepsilon = h\nu$ with Einstein's mass-energy equivalence relation, $\varepsilon = mc^2$. De Broglie reasoned that all moving particles, whose energy could be related to their mass (as expressed by Einstein's equation), also exhibited wave-like properties as described by Planck's equation (which was expressed in terms of frequency). This 'wave-particle duality' was only apparent in extremely small particles (e.g., photons and electrons) due to the small size of Planck's constant, but the effects were present in principle in macro-objects as well.

1925 saw Schrödinger develop his famous wave equations for the variation in a quantum system. The interpretation of the wave equation was a matter of some controversy, however: Schrödinger, who had realist sympathies, interpreted the wave function as a disturbance in the electromagnetic field. In other words, for Schrödinger, the elementary particles were not actually particles at all, but waves that manifested particle-like properties. Due to problems inherent in this description, however, Schrödinger's interpretation was not well received. Born, on the other hand, interpreted the wave equation as a *probability* wave.

> ... Born's way of thinking represented a marked break with classical physics. Unlike Schrödinger, who wanted to invest an element of physical reality in the wave functions, Born argued that they are actually much more abstract. He envisaged the wave function as some kind of 'probability wave', with the modulus square of the wave function representing the measurable probability for particles in specific states. In other words, he believed that the wave function represents our *knowledge* of the state of the physical object. ... Born argued that wave mechanics tells us nothing about the state of two quantum particles (such as electrons) following a collision: we can use the theory only to obtain probabilities for the various possible states resulting from the collision. ... quantum theory appeared to have no respect for the link between cause and effect.[4]

[4]Baggott 2004, 34.

In 1927, Heisenberg formulated his famous 'uncertainty principle'. He showed that it is possible to measure either the position or the momentum of an elementary particle accurately, *but not both*, and he illustrated his reasoning by means of the following thought experiment.[5] Imagine we want to observe (through a microscope) the path, described by its position and momentum, of an electron travelling through space. Now to measure the position of the electron accurately we need to use high-frequency gamma rays. Frequency, however, is directly proportional to energy via Planck's relation $\varepsilon = h\nu$. Thus as the frequency of the beam increases, so does its intensity; the energy of the individual photons in that beam becomes very high. Now there is a phenomenon known as the Compton effect: when a high energy photon collides with an electron, the electron is 'jolted' (i.e., it is knocked off its path). But what this means is that making an accurate determination of the electron's position renders us incapable of likewise accurately determining its momentum; to determine its momentum we require the position of the electron at two points along its path. But since the path has been altered by the first position measurement, we cannot determine where the electron would have been had we not interfered with it. On the other hand, if we use light of a very *low* frequency in order to avoid the Compton effect, we can then determine the electron's momentum accurately (by measuring, for example, the Doppler effect). But in that case, due to the low frequency of the beam we cannot measure its *position* accurately. Using Born's probabilistic interpretation of the wave function, Heisenberg was able to express the relationship between the uncertainties in position and momentum mathematically. He found that an *exact* determination of the position of the electron resulted in an *infinite* uncertainty in its velocity (and hence momentum), and vice versa.

[5]Cf. Baggott 2004, 36-38.

3.2 Complementarity and the Copenhagen Interpretation

While Bohr was in complete agreement with Heisenberg's conclusions, he took issue with Heisenberg's description of the experiment. Heisenberg's description assumes a particle interpretation of the electron and photon throughout, but for Bohr[6] the uncertainty arises due to the limited applicability of the particle and wave descriptions themselves to quantum phenomena. Camilleri describes Bohr's disagreement with Heisenberg as follows:

> In his detailed conceptual analysis of the gamma-ray microscope, Bohr argued that when a beam of light impinges on an electron, the light must be conceptualized both as a particle and as a wave. In the first instance it is necessary to think of the interaction between the electron and light ray as an instance of Compton scattering, according to which a photon (or a light quantum) colliding with the electron will be deflected at a given angle. However, the light quantum interpretation alone cannot explain the diffraction of the light beam, nor can it account for the resolving power of the microscope. According to classical optics, the diffracted beam of light is not scattered in a definite direction, but rather spreads over a certain angle. In the gamma-ray microscope experiment, Bohr argued, the light ray 'spreads out' like a wave after interacting with the electron, and therefore it is impossible to determine precisely at what angle the photon is scattered. ... For Bohr, it was not discontinuity of the interaction between the electron and the light quantum, but rather *the wave-particle duality of light* that prevents one from measuring the position and momentum of the electron with unlimited precision.[7]

For Bohr, the problem is not that *there is* a particle - an electron or a photon - for which we would like to determine both position and momentum but cannot. Rather, for Bohr the uncertainty arises because with respect to one type of experiment we must *assume* a particle description of the phenomenon, while with respect to another type of experiment we must *assume* a wave description - but both of these (classical) descriptions are actually incompatible with each other; we can only measure a 'particle's' momentum if we assume that it is a wave; but a wave description makes it impossible to measure the 'particle's' position. And vice versa.

[6]Heisenberg was later convinced of this as well.
[7]Camilleri 2007, 189.

> The fundamental contrast between the quantum of action and the classical concepts is immediately apparent from the simple formulae which form the common foundation of the theory of light quanta and of the wave theory of material particles. If Planck's constant be denoted by h, as is well known,
>
> $$E\tau = I\lambda = h, . \quad . \quad . \quad (1)$$
>
> where E and I are energy and momentum respectively, τ and λ the corresponding period of vibration and wave-length. In these formulae the two notions of light and also of matter enter in sharp contrast. While energy and momentum are associated with the concept of particles, and hence may be characterised according to the classical point of view by definite space-time co-ordinates, the period of vibration and wave-length refer to a plane harmonic wave train of unlimited extent in space and time.[8]

In other words, Planck's constant relates, in each case (i.e., for light and matter), two strictly speaking incompatible quantities. In the first relation, E (energy) is associated with the concept of a particle, given in time and space, while τ (period of vibration) is associated with a wave: a dynamic concept 'of unlimited extent', not conceptualizable with respect to space and time. Likewise for I and λ.

Now in the mathematical formalism of quantum mechanics (i.e., matrix and wave mechanics), when we speak of the state of a quantum *system*, this system always includes the measuring instrument as an integral part; but if the phenomenon under investigation is the 'object' and if the measuring instrument is the 'subject', then strictly speaking we cannot speak of a subject-object distinction in this context. For Bohr, this means that with respect to atomic phenomena there simply is no 'object' in the classical sense; for a distinction between the subject (represented by the measuring instrument), the object (electron, photon, etc.), and the way that the subject and object interact - which is necessary for the full determination of a thing - is impossible due to the very nature of the phenomena under investigation.

> Now the quantum postulate implies that any observation of atomic phenomena will involve an interaction with the agency of observation not to be neglected.

[8]Bohr 1928, 581.

> Accordingly, an independent reality in the ordinary physical sense can neither be ascribed to the phenomena nor to the agencies of observation. After all, the concept of observation is in so far arbitrary as it depends upon which objects are included in the system to be observed. ... This situation has far-reaching consequences. On one hand, the definition of the state of a physical system, as ordinarily understood, claims the elimination of all external disturbances. But in that case, according to the quantum postulate, any observation will be impossible, and above all, the concepts of space and time lose their immediate sense. On the other hand, if in order to make observation possible we permit certain interactions with suitable agencies of measurement, not belonging to the system, an unambiguous definition of the state of the system is naturally no longer possible, ...[9]

To illustrate this, suppose I wish to describe a particular book that is sitting on my desk (it happens to be named "Niels Bohr, Collected Works, Volume 6", and it is actually very heavy). Now there are three components involved here: first, there is the subject (myself); second, there is the object (the book); third, there is the interaction between myself and the book. We can describe this last component by means of the light rays that reflect off the book and travel through the air into my eyes. Now all three components are logically separable, and indeed, that they can be so separated is always assumed in our descriptions of physical phenomena. Now suppose, for example, that the book is not actually on my desk, but underwater, and I am standing on the shore. In this case the book will appear to be displaced from its actual position due to the refraction of the light wave through a second medium (the water). Now I am able to make the necessary corrections in this case. I know about the laws for the refraction of light (or at least I am aware of them), and taking these into account, I am able to determine, for instance, the actual position of the book with reasonable certainty. In both cases, in order to give a precise description of this book, it was necessary to separate and describe the three components: subject, object, and their interaction. In the first case we needed to do this in order to show that the book actually is where it appears to be, and in the second case, to show that it is not. In quantum mechanics, however, it is *impossible to actually describe the interaction* between

[9]Bohr 1928, 580.

the measuring instrument and the electron.[10] It is, of course, possible to describe the overall quantum system using wave or matrix mechanics - but this description of the system is not one in which subject and object can be distinguished.

All physical description presupposes that there is something to be described; it presupposes an interaction between subject and object in which all three aspects can be distinguished. But if such a distinction is impossible in quantum mechanics, then strictly speaking the 'classical' concepts: space, time, and causality are *inapplicable.* But we nonetheless *require* these concepts, for Bohr, in order to conceptualize our experiments; we need them to interpret our experience and to communicate this experience; they are, in a certain sense, 'a priori' concepts.

> ... the requirement of communicability of the circumstances and results of experiments implies that we can speak of well defined experiences only within the framework of ordinary concepts. In particular it should not be forgotten that the concept of causality underlies the very interpretation of each result of experiment, and that even in the coördination of experience one can never, in the nature of things, have to do with well-defined breaks in the causal chain. The renunciation of the ideal of causality in atomic physics which has been forced on us is founded logically only on our not being any longer in a position to speak of the autonomous behavior of a physical object, due to the unavoidable interaction between the object and the measuring instruments which in principle cannot be taken into account, if these instruments according to their purpose shall allow the unambiguous use of the concepts necessary for the description of experience.[11]

In order to describe the electron at all we must *presuppose* that the electron is an object to be described, i.e., that it is something that can be determined with regard to space and time and with regard to its interaction with other objects. For example, if an electron leaves a mark on a photographic plate, then we must interpret this as a mark left by an *object* that interacted with the plate in a specific area, i.e., in space and time. Similarly, the interaction between the measuring instrument and the object must itself be described *causally* -

[10]Cf. Bohr 1928, 585, 586.
[11]Bohr 1937, 293.

as a *transfer* of energy and momentum.[12] Now the uncertainty relations tell us that neither of these (position and momentum) can be determined *accurately* without foregoing a precise determination of the other. Thus it is *impossible in principle* to obtain a *determinate* conception of atomic phenomena. But ironically, it is the uncertainty relations themselves which enable us to go about our merry way and continue doing physics. For by means of the uncertainty relations, we can act in certain circumstances 'as if' the phenomena under investigation were a wave: when we wish to determine its momentum, and in certain circumstances 'as if' it were a particle: when we wish to determine its position. In reality, however, if the phenomena under investigation *were* a particle, *it could not* be also a wave, and vice versa. The two descriptions are mutually contradictory. But since we *cannot* determine the electron fully, apart from its interaction with the measuring apparatus, the electron *is not* a determinate object and therefore we are justified in using *either* description of it; both descriptions can be thought of as *complementary*; for they both serve us in our effort to interpret and communicate the results of different experiments. These descriptions can be thought of as abstractions, or idealizations - not strictly speaking applicable, but usable and necessary.

> From the above considerations it should be clear that the whole situation in atomic physics deprives of all meaning such inherent attributes as the idealizations of classical physics would ascribe to the object. On the contrary, the proper rôle of the indeterminacy relations consists in assuring quantitatively the logical compatibility of apparently contradictory laws which appear when we use two different experimental arrangements, of which only one permits an unambiguous use of the concept of position, while only the other permits the application of the concept of momentum ...[13]

But coming back to the question of scientific explanation: if Bohr's view is that it is *necessary* to speak in the language of classical physics, i.e., in terms of space, time, and causality, then isn't this just a vindication of the view that I have been arguing against?

[12] See also: Camilleri 2005, 274.

[13] Bohr 1937, 293. See also Bohr 1928, 581.

No, this is not the case; for first, on that view the causal laws must be *true* laws: they must describe physically real interactions between *objectively* real entities - whether we consider these entities to be point-like particles or causal processes. But since, on the Copenhagen interpretation it is strictly speaking impossible to speak of the *object* of investigation, the causal interactions we describe are not strictly speaking true, since they are based on a strictly speaking *false* description of what we observe. Second, with regard to position measurements, since we cannot determine the momentum of the 'object', we *cannot* describe this 'object', even falsely, in terms of its causal interaction with other objects at all. For a determination of the 'particle' in terms of space and time *foregoes* a determination of it in terms of its causal interactions. Thus with regard to the description of the results of certain experiments an appeal to causes is strictly speaking false (though still useful and moreover necessary), and with regard to certain other experiments it is absent entirely.

> Indeed this circumstance presents us with a situation concerning the analysis and synthesis of experience which is entirely new in physics and forces us to replace the ideal of causality by a more general viewpoint usually termed "complementarity." The apparently incompatible sorts of information about the behaviour of the object under examination which we get by different experimental arrangements can clearly not be brought into connection with each other in the usual way, but may, as equally essential for an exhaustive account of all experience, be regarded as "complementary" to each other.[14]

But one may wonder, now, what all the fuss is about; i.e., one may wonder why we should view ourselves as limited by these classical concepts. Perhaps quantum mechanics requires a new conceptual framework altogether. Cassirer, for example, expresses this sentiment when he writes:

> We must face squarely the new problems thus created. There seems to be no return to the lost paradise of classical concepts; physics has to undertake the construction of a new methodological path. I do not wish to claim at all that the end of this path is already clearly in sight. But the direction in which the

[14]Bohr 1937, 291.

> solution is to be found seems to me clearly recognizable. Physics and epistemology cannot continue to posit a being with full realization that it contradicts the conditions of physical knowledge. If it appears that certain concepts, such as those of position, of velocity, or of the mass of an individual electron can no longer be filled with a definite empirical content, we have to exclude them from the theoretical system of physics, important and fruitful though their function may have been.[15]

Indeed, it certainly is possible to describe atomic phenomena by means of the mathematical formalism of either wave or matrix mechanics. Why must we insist on hanging on to the old language? Bohr points out, however, that the situation is analogous to that in the special theory of relativity: "the fundamental constants, the velocity of light and the quantum of action, are introduced into the formalism as factors of the $\sqrt{-1}$, the one in the definition of the fourth coördinate, the other in the commutation laws of canonically conjugate variables. ... in these fields the logical correlations can only be won by a far-reaching renunciation of the usual demands of visualization."[16] For Bohr, while the mathematical formalism is critical for an *analysis* of our experience in the atomic domain, the *synthesis* of that experience: the connection of different experiences by which we comprehend and communicate these experiences, requires that we use ordinary concepts.[17]

3.3 Kantian Aspects of Complementarity

Anyone familiar with Kant's philosophy will have immediately recognized some striking similarities in the arguments of these two thinkers. Kant makes use of something analogous to Bohr's 'principle of complementarity' often in his own philosophy. It is, in fact, his doctrine of the antinomies: a method for showing that two seemingly contradictory arguments are either both false or both compatible. In the latter case, Kant's resolution always consists in showing that the thesis and antithesis of the antinomy, though contradic-

[15]Cassirer [1936] 1956, 194-195.
[16]Bohr 1937, 292.
[17]Cf.Bohr 1937. 292-293.

tory when considered as pertaining to things *as they are in themselves*, are compatible if we consider them as applying merely to things as they appear to us.

> If one regards the two propositions, that the world is infinite in magnitude and that the world is finite in magnitude, as opposed to each other contradictorily, then one assumes that the world (the entire series of appearances) is a thing in itself. For in either proposition the world remains, whether I annul in the series of its appearances the infinite or the finite regression. But if I remove this presupposition - or i.e. this transcendental illusion - and deny that the world is a thing in itself, then the contradictory conflict of the two assertions is transformed into a merely dialectical one ...[18]
>
> Such a twofold side from which to think the power of an object of the senses contradicts none of the concepts that we have to frame of appearances and of a possible experience. For since these appearances are not in themselves things, they must be based on a transcendental object determining them as mere presentations; and hence nothing prevents us from attributing to this transcendental object, besides the property through which it appears, also a *causality* that is not appearance although its *effect* is nonetheless encountered in appearance.[19]
>
> Now, to show that this antinomy rests on a mere illusion and that nature at least does *not conflict* with the causality from freedom - this was the only goal that we were able to accomplish, and it was, moreover, our one and only concern.[20]

Now it is certainly not my intention to affix to Bohr the label of 'Kantian'; Bohr's philosophy was in many ways completely unique. Moreover, Bohr did not actually explicitly refer to Kant very much, either in his published or unpublished writings. On the other hand, his education did occur at a time in which Neo-Kantianism in general was in the ascendency. But as interesting as this question is, it is not my intention to engage in an historical analysis of the extent to which Bohr was or was not directly or indirectly influenced by Kant in his overall thinking. While such an historical analysis certainly has its merits,[21] my aim is only to show that there are similarities with respect to the central doctrines of

[18]Kant [1781/1787] 1996, B532-533.

[19]Kant [1781/1787] 1996, B566-567.

[20]Kant [1781/1787] 1996, 586.

[21]For some points of view on this matter, see, for example: Folse 1978, Howard 2004, Honner 1982, Camilleri 2005, Camilleri 2007.

these two thinkers which can help us to understand the both of them better, and which are relevant to the explanation of atomic phenomena.

It is often said that quantum mechanics, because it abrogates the law of causality, or because it entails irreducibly probabilistic phenomena, refutes (or at least contradicts) Kant. It does not do so, however - at least not in essentials - if we follow the Copenhagen interpretation. Kant distinguishes between *mathematical* and *dynamical* principles for the possibility of experience.[22] The former are *constitutive for appearances*. They are necessary principles for the possibility of presenting an appearance to ourselves as *existing*. According to these principles, in order for anything to appear to us, it must be apprehended as having both an extensive (length, breadth, etc.) and an intensive magnitude (i.e., a degree). The *dynamical* principles, on the other hand, are not constitutive but merely *regulative*.[23] They are principles not for the apprehension, but for the *connection* of appearances in time; therefore they *presuppose that an appearance already exists* (i.e., that it has already been apprehended in accordance with the mathematical principles). These dynamical principles are, first, that all change presupposes something that is permanent (in the language of physics, we might call this the principle of conservation of matter). Second, all changes must occur according to the law of cause and effect. Third, all substances that are perceived as simultaneous are in mutual interaction.

Now since the dynamical principles are merely regulative (i.e., since they are only valid with regard to actually existing appearances), that something appears must be presupposed in order for a time determination of it according to the dynamical principles to be possible. But that something appears to us as existing is, by itself, not enough to determine this something as an *object*. To *determine* this appearance as an object, we must apply the dynamical principles to it. We must say that, at a determinate instant of time, it has a deter-

[22] Kant [1781/1787] 1996, B198-B294.

[23] This sense of regulative should not be confused with the sense that Kant uses with respect to the 'ideas of reason'. There the distinction is between that which is constitutive or regulative with respect to experience as a whole. Here, he uses regulative not in the context of experience in general, but in the context of particular objects of experience.

minate position in space (determined by the mathematical principles) and that there is a law (subject to the dynamical principles) by which it dynamically interacts with other objects. Now consider atomic phenomena. By means of the mark, say, on a photographic plate, we are able to determine that something, which we conceptualize as a particle, exists, and that it interacted with the plate at such-and-such a place - but by means of this experiment we cannot in principle determine it with regard to its momentum; in other words we cannot determine its dynamical interaction with the plate. Similarly we can, by means of another experiment, determine the momentum of this something if we conceptualize it as a wave that is unlimited in extent; but by doing so we forego any determination of this something as existing determinately at a definite position. Since atomic phenomena cannot be simultaneously determined with regard to the mathematical and dynamical principles, both of which are required for the determination of an *object,* atomic phenomena simply are not objects in the everyday sense of the term, and both the wave and particle descriptions are, strictly speaking, *false* if we take them as being applicable to atomic phenomena as they are *in themselves*. It is no surprise at all, then, that according to the precise mathematical formalism of quantum mechanics, we are unable to distinguish object from measuring instrument - for there is no object to distinguish - not, that is, in the sense that we are capable of understanding something as an object. And it is no surprise at all that the principle of causality does not apply to these phenomena.

In Kant's language, we would call these 'objects' *noumena*, and it is because they are noumena that we can apply abstractions such as a wave or a particle description to them without fear of contradiction, so long as we do not understand these descriptions to be applicable to them as they are in themselves. To clarify: according to Kant, a concept of the understanding must be understood both in terms of its form and in terms of the content to which it can be applied. We can think of the form of a concept as analogous to a mathematical function, e.g., $f(x) = 2x + 4$. Now a determinate result can be obtained for this function only if something is filled in for x. By itself, the function only represents a form for the determination of a variable. Likewise for a concept: without content, a concept gives us no *determinate* cognition. "Without an object the concept has no sense and is completely empty of content, although it may still contain the logical function for making a concept from what data may come up".[24]

For the case of the mathematical function, the only allowable values of x (i.e., the domain of x) are numbers. Thus, $f(5)$ is defined but $f(\text{'}A\text{'})$ is not. Now the concept of a noumenon (in the negative sense of the term) is the concept of something *indeterminate*. It is analogous to x in the mathematical equation. The function above *cannot* be applied to x itself, but only to a value that has been filled in for x. Similarly for concepts: they cannot be applied to an indeterminate intuition but only to a determinate one. But now consider atomic phenomena. Because we are unable - in principle according to the Copenhagen interpretation - to make a distinction between the object being measured and the instrument doing the measuring, we cannot represent this object determinately as existing outside 'ourselves' (i.e., the measuring instrument). Therefore we cannot think of the appearance of the elementary particle as the appearance of a determinate something *existing*, in the strict sense of that word, and therefore the dynamical principles: permanence, causality, simultaneity, etc. - those principles stemming from the concepts of our understanding which determine the appearance in time - simply *cannot be applied* to it. Atomic phenomena are *problematic concepts*:

> I call a concept problematic if, although containing no contradiction and also cohering with other cognitions as a boundary of given concepts involved in them, its objective reality cannot be cognized in any way. The concept of *noumenon*, i.e., of a thing that is not to be thought at all as an object of the senses but is to be thought (solely through a pure understanding) as a thing in itself, is not at all contradictory; for we cannot, after all, assert of sensibility that it is the only possible kind of intuition. Moreover, the concept of a noumenon is necessary in order not to extend sensible intuition even over things in themselves, and hence in order to *limit the objective validity of sensi-*

[24] Kant [1781/1787] 1996, B298.

ble cognition.[25]

Thus if one considers an electron as it is in itself - a concept for which we can have no sensible intuition - then the wave and particle descriptions are indeed contradictory. But if one considers only the appearances resulting from our experiments - and these are given in space and time - then these descriptions are compatible, i.e., they are *complementary* descriptions of what we observe.

Now Kant held these 'classical' concepts to be *absolutely* a priori - necessary concepts for a rational being as such. I do not wish to argue for or against this view here, but I will note that there appears to be nothing barring us from 'relativizing' Kant's conception of the a priori. It is not clear what Bohr's opinions on this were, however this was Heisenberg's view: "... the Kantian 'a priori' is indirectly connected with experience in so far as it has been formed through the development of the human mind in a very distant past. ... It is in fact quite plausible that for certain primitive animals space and time are different from what Kant calls our 'pure intuitions' of space and time."[26] And also: "We should nevertheless remember that the very structure of human thought changes in the course of historical development. Science progresses not only because it helps to explain newly discovered facts, but also because it teaches us over and over again what the word 'understanding' may mean."[27]

The Copenhagen interpretation of quantum mechanics is certainly not the only available interpretation of atomic phenomena on offer. Although it was for a long time, the *only* interpretation taught to physics students, in recent years additional interpretations have been proposed. Two such examples are the de-Broglie-Bohm theory and Popper's propensity theory.[28] Now I do not intend to engage in a detailed discussion of the relative merits and shortcomings of each of these interpretations. It is enough, for my limited purposes, to

[25]Kant [1781/1787] 1996, B310. Emphasis added.
[26]Heisenberg 1959, 83.
[27]Heisenberg 1971, 124.
[28]For some of the criticisms of the Copenhagen interpretation and some alternative interpretations, see: Einstein et. al. 1935, Popper 1982, Bunge 1955, Bunge 1955a, Bohm 1957, Baggott 2004. For Bohr's response to Einstein, see: Bohr 1935 and Bohr 1949.

point out that the Copenhagen interpretation presents us with a non-causal explanation of atomic phenomena, and importantly: quantum mechanics is not some trivial exception to the rule of causal explanations in science; it is the most successful theory in physics.

3.4 Probabilistic Causality, Causal Processes, and Reductionism

Let us return now to some of the questions that we left open earlier in this thesis, and consider whether anything we have discussed in this chapter can help to shed some light on them. Before we do, however, let me reiterate that my intention was clearly not to present an argument *for* the orthodox Copenhagen interpretation of quantum mechanics. I have not provided anything like a defence of this interpretation against the criticisms levelled at it by some. Rather, I have simply taken it for granted that this theory which for close to a century has enjoyed such widespread acceptance, accounted for every experimental result, and withstood the many thought experiments and actual experiments designed to refute it, can tell us something about the subject of scientific explanation. It is, after all, the most successful theory in physics. Philosophy is not a popularity contest, however, and we should never appeal to consensus when we can appeal to sound argument instead. Thus I will leave it open whether or not the Copenhagen interpretation is actually correct. Assuming that it is, however (and the burden of proof is not on it), what can quantum mechanics tell us about the questions of probabilistic causality, causal processes, and reductionism, as they relate to the question of scientific explanation?

Let us begin with probabilistic causes, which we discussed in chapter 2. First, I think that we need to distinguish between probabilistic processes that are reducible in principle to deterministic ones, and those that are not. With regard to the latter we can say that these all belong to the domain of atomic physics. This, I take it, is uncontroversial, and according to the dominant interpretation of atomic physics the principle of causality cannot strictly speaking be applied to these processes at all. Thus if we accept the Copenhagen interpretation (as Railton purports to do), then there is no need to appeal to propensities in order to interpret these probabilities causally, indeed such an interpretation of probability would seem to actually contradict the Copenhagen interpretation of quantum mechanics.

As for probabilistic processes that are reducible in principle to deterministic processes, I think it indeed makes sense to speak of 'probabilistic causes' in these cases - so long as we only consider probability to be a good *indicator* of an underlying deterministic mechanism and do not view these probabilities as *themselves* explicating the notion of a causal link.[29] Recall that on the model of explanation that I have been defending (the D-N-P), the requirement is that we always explain *reducibly* probabilistic processes deterministically. Thus an ideal explanatory text for an apparently probabilistic process would not contain any probabilistic laws at all. We could still appeal to reducibly probabilistic laws in an explanation, however, and on the D-N-P model, via Railton's notion of explanatory information, such an explanation would be perfectly valid so long as these laws are seen merely as approximations.

Let us now turn to the characterization of causal processes themselves, and specifically to the debate between Salmon and Dowe. Recall that Dowe characterized a causal process as the world line of an object which possesses a conserved quantity.[30] Salmon, recall, criticized Dowe on the grounds that some objects (e.g., the portion of a wall illuminated by a spotlight as the spotlight moves back and forth) satisfy Dowe's conserved quantity criterion but nevertheless do not qualify as causal processes because they cannot transmit energy. Dowe responded that such an object as the illuminated spot on a wall cannot qualify as an object for it is not something that wholly exists *in* time, to which Salmon responded that 'wholly existing at a time' begs the question of the *explication* of genidentity (identity over time).

Now if we believe Bohr then we should side with Dowe here. For according to Bohr,

[29]On the irreducibility of causes to probabilities, see Cartwright 1989, 22-38.
[30]Cf. Above.

the concept of causality is simply not applicable unless we can distinguish an *object*; and an object, on this view, is exactly what Dowe describes it to be: something that is 'wholly present at a time'. As for the issue of genidentity, this is a *non sequitur*. An object is whatever we, for pragmatic reasons, take it to be - just so long as *whatever we define it to be* can be determined both in and across time. In quantum mechanics also, the line that we draw between measuring instrument and 'object' is always arbitrary: we must always draw this line if we are to communicate our experiences of atomic phenomena, but we always draw it so that it is most convenient for our purposes.

As for the question of reductionism which we discussed at the end of chapter 1, I think the answer is clear that according the Copenhagen interpretation the reductionist thesis should be rejected. According to Bohr's principle of complementarity, the 'object' of investigation must, as we have seen, *necessarily* be characterized according to two mutually incompatible but complementary descriptions. Further, he writes:

> Just the fact that the paradoxes of atomic physics could be solved not by a one sided attitude towards the old problem of "determinism or indeterminism," but only by examining the possibilities of observation and definition, should rather stimulate us to a renewed examination of the position in this respect in the biological and psychological problems at issue. ... The only logical possibility of avoiding any contradiction between the formulation of the laws of physics and the concepts suitable for the description of the phenomena of life ought therefore to be sought in the essentially different character of the conditions of investigation concerned.[31]

Bohr never pursued this line of thought, unfortunately; however his sentiments here are strikingly similar to Kant's, and Kant indeed has both an ethical and a biological theory *grounded* on something very similar to Bohr's principle of complementarity. If Bohr's principle of complementarity and the Copenhagen interpretation of quantum mechanics can be defended it is but one more argument for a non-reductionist, Kantian-inspired view of ethics and biology. If Bohr and Kant are right, then we are completely justified in

[31] Bohr 1937, 295.

denying to Railton that there can be *only one* ideal text for any event, and instead holding that can be many, perhaps as many as one per special science.

But suppose the Copenhagen interpretation of quantum mechanics is incorrect (recall that I have not defended it). In this case we should not be worried, for the model of explanation that I have defended is perfectly compatible with a propensity interpretation of probability; it is perfectly compatible with a reductionist view of science, and it is perfectly compatible with a causal interpretation of irreducibly probabilistic processes. But of course it is, for these are Railton's own views.

Conclusion

Railton's D-N-P model begins with Hempel's D-N model but has the advantage that it makes room for - indeed, it stresses - some of our intuitions concerning causal cases. Further, his distinction between an ideal explanatory text and explanatory information allows us to take pragmatic considerations into account when evaluating explanations. The requirement to give a full account of the mechanism giving rise to the explanandum is enough to deal with both the irrelevance and asymmetry counter-examples. Railton's D-N-P model is, I think, the most promising model of singular explanation yet put forward. Yet an account of explanation should make as few metaphysical commitments as possible. A commitment to hard reductionism or to the propensity interpretation of probability has no place, and I believe that I have shown how Railton's D-N-P model can function perfectly well without them.

The modified version of the D-N-P model that I have defended in this thesis is neutral with respect to the issue of reductionism, with respect to the interpretation of probability, and with respect to a causal characterization of irreducibly probabilistic processes. In chapter 1, I argued that Railton's requirement that there be one and only one ideal explanatory text for an event should be dropped; in chapter 2, I argued that there was no need to appeal to a propensity interpretation of probability in order to account for irreducibly probabilistic processes; and throughout, I have been arguing that an account of scientific explanation that focuses solely on causes is incomplete. In chapter 3, finally, I showed how the denial of all of these views is compatible with our modern physical theories.

Kolmogorov's highly successful axiomatization of the probability calculus is, for me, an excellent example of the type of approach that I believe is best suited for an account of explanation. Although Kolmogorov's own views on probability are frequentist, his axiomatization makes no assumptions and takes no interpretation of prior probabilities for granted. Its success is a witness to what can be achieved when we focus on the issue at hand. An account of explanation should not attempt to solve all the problems extant in the philosophy of science - only the ones with which we are concerned.

Bibliography

[1] Achinstein, Peter. *The Nature of Explanation.* New York: Oxford University Press, 1983.

[2] Baggott, Jim. *Beyond Measure: Modern Physics, Philosophy and the Meaning of Quantum Theory.* New York: Oxford University Press, 2004.

[3] Black, Max. *A Companion to Wittgenstein's Tractatus.* Ithaca, NY: Cornell University Press, 1966.

[4] Bohm, David. *Causality and Chance in Modern Physics.* Philadelphia: University of Pennsylvania Press, 1957.

[5] Bohr, Niels. "The quantum postulate and the recent development of atomic theory." *Nature* 121 (1928): 580-590.

[6] Bohr, Niels. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review* 48 (1935): 696-702.

[7] Bohr, Niels. "Causality and Complementarity." *Philosophy of Science* 4.3 (1937): 289-298.

[8] Bohr, Niels. "Discussion with Einstein on Epistemological Problems in Atomic Physics." In *Albert Einstein: Philosopher-Scientist,* edited by Paul Arthur Schlipp, 201-241. La Salle, IL: Open Court, 1949.

[9] Bromberger, Sylvain. "Why Questions." In *Mind and Cosmos,* edited by R.G. Colodny, 86-108. Pittsburgh: University of Pittsburgh Press, 1966.

[10] Bunge, Mario. "Strife about Complementarity (I)." *The British Journal for the Philosophy of Science.* 6.21 (1955): 1-12.

[11] Bunge, Mario. "Strife about Complementarity (II)." *The British Journal for the Philosophy of Science.* 6.22 (1955): 141-154.

[12] Camilleri, Kristian. "Heisenberg and the Transformation of Kantian Philosophy." *International Studies in the Philosophy of Science.* 19.3 (2005): 271-287.

[13] Camilleri, Kristian. "Indeterminacy and the Limits of Classical Concepts: The Transformation of Heisenberg's Thought." *Perspectives on Science.* 15.2 (2007): 178-201.

[14] Carnap, Rudolf. *Logical Foundations of Probability*. Chicago: University of Chicago Press, 1962.

[15] Cartwright, Nancy. *Nature's Capacities and their Measurement*. Oxford: Clarendon Press, 1989.

[16] Cassirer, Ernst. *Determinism and Indeterminism in Modern Physics*. Translated by O. Theodor Benfey. New Haven: Yale University Press, 1956.

[17] Curd, Martin, and J.A. Cover. "Models of Explanation (Commentary)." In *Philosophy of Science: The Central Issues*, edited by Martin Curd and J.A. Cover, 767-804. New York: W.W. Norton & Company, Inc., 1998.

[18] Dowe, Phil. "Causality and Conserved Quantities: A Reply to Salmon." *Philosophy of Science* 62 (1995): 321-333.

[19] Einstein, A., B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review* 47 (1935): 777-780.

[20] Feynman, R. et al. *The Feynman Lectures on Physics, Volume III*. Reading, MA: Addison-Wesley, 1965.

[21] Folse, Henrey J. "Kantian Aspects of Complementarity." *Kant-Studien* 69 (1978): 58-66.

[22] Friedman, Michael. "Explanation and Scientific Understanding." *The Journal of Philosophy* 71 (1974): 5-19.

[23] Galavotti, Maria Carla. *Philosophical Introduction to Probability*. Stanford: CSLI Publications, 2005.

[24] Heisenberg, Werner. *Physics and Philosophy: The Revolution in Modern Science*. London: George Allen & Unwin Ltd., 1959.

[25] Heisenberg, Werner. *Physics and Beyond*. Translated by Arnold J. Pomerans. New York: Harper & Row, 1971.

[26] Hempel, Carl G. *Aspects of Scientific Explanation And Other Essays in the Philosophy of Science*. New York: The Free Press, 1965.

[27] Hempel, Carl G. "Laws and Their Role in Scientific Explanation." In *Philosophy of Natural Science*, 47-69. Englewood Cliffs, NJ: Prentice-Hall Inc., 1966.

[28] Honner, John. "The Transcendental Philosophy of Niels Bohr." *Studies in History and Philosophy of Science* 13 (1982): 1-29.

[29] Howard, Don. "Who Invented the "Copenhagen Interpretation"? A Study in Mythology." *Philosophy of Science* 71 (2004): 669-682.

[30] Humphreys, Paul. "Why Propensities Cannot be Probabilities." *The Philosophical Review* 94.4 (1985): 557-570.

[31] Humphreys, Paul. "Scientific Explanation: The Causes, Some of the Causes, and Nothing But the Causes." In *Scientific Explanation (Minnesota Studies in the Philosophy of Science, Volume XIII)*, edited by Philip Kitcher and Wesley C. Salmon, 283-306. Minneapolis: University of Minnesota Press, 1989.

[32] Kitcher, Philip. "Explanatory Unification and the Causal Structure of the World." In *Scientific Explanation (Minnesota Studies in the Philosophy of Science, Volume XIII)*, edited by Philip Kitcher and Wesley C. Salmon, 410-505. Minneapolis: University of Minnesota Press, 1989.

[33] Kant, Immanuel. *Critique of Pure Reason*. Translated by Werner S. Pluhar. Indianapolis: Hackett Publishing Company, 1996.

[34] Laplace, Pierre Simon. *Essai philosophique sur les probabilités*. Paris: Courcier, 1825.

[35] McGuinness, Brian, ed. Schulte, Joachim and Brian McGuinness, trans., *Wittgenstein and the Vienna Circle: conversations recorded by Friedrich Waismann* (Oxford: Basil Blackwell, 1979),

[36] Popper, Karl R. "The Propensity Interpretation of Probability." *The British Journal for the Philosophy of Science* 10.37 (1959): 25-42.

[37] Popper, Karl R. *Quantum Theory and the Schism in Physics*. Totowa, NJ: Rowman and Littlefield, 1982.

[38] Railton, Peter. "A Deductive-Nomological Model of Probabilistic Explanation." *Philosophy of Science* 45 (1978): 206-226.

[39] Railton, Peter. "Probability, Explanation, and Information." *Synthese* 48 (1981): 233-256.

[40] Reichenbach, Hans. *The Theory of Probability: An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability*. tr. Ernest H. Hutten and Maria Reichenbach. Berkely and Los Angeles: University of California Press, 1949.

[41] Ruben, David-Hillel. "Arguments, Laws, and Explanation." In *Philosophy of Science: The Central Issues*, edited by Martin Curd and J.A. Cover, 720-745. New York: W.W. Norton & Company, Inc., 1998.

[42] Salmon, Wesley C. *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press, 1984.

[43] Salmon, Wesley C. "Four Decades of Scientific Explanation." In *Scientific Explanation (Minnesota Studies in the Philosophy of Science, Volume XIII)*, edited by Philip Kitcher and Wesley C. Salmon, 3-219. Minneapolis: University of Minnesota Press, 1989.

[44] Salmon, Wesley C. "Causality Without Counterfactuals." In *Philosophy of Science* 61 (1994): 297-312.

[45] Salmon, Wesley C. "Causality and Explanation: A Reply to Two Critiques." In *Philosophy of Science* 64 (1997): 461-477.

[46] Scriven, Michael. "Explanations, Predictions, and Laws." In *Scientific Explanation, Space, and Time (Minnesota Studies in the Philosophy of Science, Volume III)*, edited by Herbert Feigl and Grover Maxwell, 170-230. Minneapolis: University of Minnesota Press, 1962.

[47] van Fraassen, Bas C. *The Scientific Image*. Oxford: Clarendon Press, 1980.

[48] von Mises, Richard. *Probability, Statistics and Truth*, tr. J. Neymann, D. Scholl and E. Rabinowitsch. London: George Allen and Unwin Ltd., 1957.

[49] von Wright, G.H. "Wittgenstein's Views on Probability", *Revue Internationale de Philosophie* 23 (1969): 259-283.

[50] Wittgenstein, Ludwig. *Philosophical Remarks*, tr. Raymond Hargreaves and Roger White, ed. Rush Rhees. Oxford: Basil Blackwell, 1975.

[51] Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*, tr. D.F. Pears and B.F. McGuinness. London: Routledge, 2005.