

CAUSAL EXPLANATION AND SCIENTIFIC REALISM

ABSTRACT: It is widely believed that many of the competing accounts of scientific explanation have ramifications which are relevant to the scientific realism debate. I claim that the two issues are orthogonal. For definiteness, I consider Cartwright's argument that causal explanations secure belief in theoretical entities. In Section I, van Fraassen's anti-realism is reviewed; I argue that this anti-realism is, *prima facie*, consistent with a causal account of explanation. Section II reviews Cartwright's arguments. In Section III, it is argued that causal explanations do not license the sort of inferences to theoretical entities that would embarrass the anti-realist. Section IV examines the epistemic commitments involved in accepting a causal explanation. Section V presents my conclusions: *contra* Cartwright, the anti-realist may incorporate a causal account of explanation into his vision of science in an entirely natural way.

It is a widely held myth that one's position on the issue of scientific realism puts constraints on the range of theories of scientific explanation that one can consistently maintain. Peter Railton, for example, writes:

What else should one expect? It is inconceivable that a notion such as explanation could fail to depend crucially upon one's most general picture of the world and its ways. (1989, p. 221)

I wish to argue that a traditionally 'realist' account of explanation—the causal account—is compatible with the sort of anti-realism developed in van Fraassen (1980).

The principal link between the scientific realism debate and competing theories of scientific explanation is provided by a family of inference patterns which march under the slogan 'inference to the best explanation'.¹ In its simplest form, the inference pattern is

$$\frac{Q}{P \text{ explains } Q} \\ p^2$$

Many (see, e.g., Glymour (1984)) have argued that an inference pattern such as this allows us to infer to the existence of unobservable entities. It would seem then, that the anti-realist is committed to a conception of explanation which invalidates this inference pattern. Van Fraassen

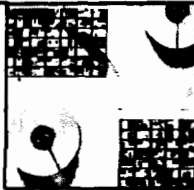
Erkenntnis 37: 151–178, 1992.

© 1992 Kluwer Academic Publishers. Printed in the Netherlands.

This publication
is available in
microform
from UMI.

U·M·I

800-521-0600 toll-free
313-761-4700 collect from Alaska and Michigan
800-343-5299 toll-free from Canada



Copies of articles from
this publication are now
available from UMI
Article Clearinghouse.

U·M·I

A Bell & Howell Company
300 North Zeeb Road
Ann Arbor, MI 48106 USA

Photocopying. In the U.S.A.: This journal is registered at the Copyright Clearance Center, Inc., 27 Congress Street, Salem, MA 01970.

Authorisation to photocopy items for internal or personal use, or the internal or personal use of specific clients is granted by Kluwer Academic Publishers for users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the fee of \$5.00 per copy is paid directly to CCC. For those organisations that have been granted a photocopy licence by CCC, a separate system of payment has been arranged. The fee code for users of the Transactional Reporting Service is 0165-0106/92\$ 5.00.

Authorisation does not extend to other kinds of copying, such as that for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale.

In the rest of the world: Permission to photocopy must be obtained from the copyright owner. Please apply to Kluwer Academic Publishers, P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Erkenntnis is published bimonthly.

Subscription price, per volume: \$ 141.50, including postage and handling.

Second-class postage paid at Rahway, N.J. USPS No. 428-770.

U.S. Mailing Agent: Expeditors of the Printed Word Ltd., 2323 Randolph Ave., Avenel, NJ 07001.

Published by Kluwer Academic Publishers, Spuiboulevard 50, P.O. Box 17, 3300 AA Dordrecht, The Netherlands, and 101 Philip Drive, Norwell, MA 02061, U.S.A.

Postmaster: Please send all address corrections to: *Erkenntnis*, c/o Expeditors of the Printed Word Ltd., 2323 Randolph Ave., Avenel, NJ 07001, U.S.A.

Printed on acid-free paper

has sharply criticized inference to the best explanation. (See, e.g., his (1980), chapter 2, §3; and (1989), chapter 6.) Nancy Cartwright has argued that inference to the best explanation is valid in the case of causal explanations:

Van Fraassen asks, what has explanatory power to do with truth? He offers more a challenge than an argument: show exactly what about the explanatory relationship tends to guarantee that if *x* explains *y* and *y* is true, then *x* should be true as well. This challenge has an answer in the case of causal explanation . . . (Cartwright 1983, p. 4)

Cartwright, although an anti-realist about theoretical laws, provides a clear and intuitive argument connecting causal explanations with realism about unobservable entities. I suspect that in this regard she speaks for many realists who have championed causal theories of explanation. I will argue that van Fraassen's challenge is not met; the bump in the rug is pushed down only to reappear elsewhere. Moreover, the anti-realist can defend against Cartwright's argument using only the weapons he has already stocked in his arsenal.

In Section I, I will sketch the opposing positions in the realism debate and motivate my claim that anti-realism is consistent with a causal conception of explanation. In Section II, I will review Cartwright's distinction between theoretical and causal explanations, and her arguments connecting causal explanations with realism about theoretical entities. In Section III, I will argue that inference to the best causal explanation does not provide the realist with an inroad to the realm of the unobservable. Section IV will examine the broader claim that acceptance of a causal explanation commits one to belief in the entities postulated in the explanation. Section V will contain my conclusions.

I

The issue of scientific realism has given rise to more lively debate over the last decade than any other issue in the philosophy of science. Bas van Fraassen, the leading opponent of scientific realism, has characterized the position he opposes as follows:

Science aims to give us, in its theories, a literally true story of what the world is like; and acceptance of a scientific theory involves the belief that it is true. (van Fraassen 1980, p. 8)

There are two components to the realist's position as characterized by van Fraassen: the first concerns the interpretation of the claims made

within a scientific theory, the second concerns our epistemic attitude toward these claims. According to scientific realism, we should interpret the claims of a scientific theory literally; and the epistemic commitment involved in accepting a theory is a strong one: belief that the theory is true. Van Fraassen agrees with the first part of the realist's position, but disagrees with the second. His position on the epistemic attitude appropriate to theory acceptance is more permissive: the observational evidence which would compel one to accept a theory does not compel one to believe what the theory says about the unobservable. Van Fraassen prefers to remain agnostic about unobservables, although he would not condemn one who chose to believe in them; he would, however, regard this belief as one *permitted*, rather than required, by the evidence. (See van Fraassen 1984, 1985, 1989.) Van Fraassen dubs his brand of anti-realism 'constructive empiricism'.

Many realists have tried to counter van Fraassen's challenge by producing inference patterns in which observational evidence does in fact compel belief in unobservables. One family of such patterns goes by the name 'Inference to the Best Explanation'. For example, the hypothesis that gas is composed of molecules in motion, together with Newton's laws of motion and the principle that temperature is a measure of mean kinetic energy, allows one to derive the Boyle-Charles gas law. Since this law is true (approximately), and the molecular hypothesis, if true, would provide an explanation of the Boyle-Charles law, we may infer that the molecular hypothesis is true. Van Fraassen has raised a number of objections to inferences such as this: that their naturalistic justification is based on inconclusive evidence (1980, chapter 2, §3); that there is no reason to expect the true explanation to be among the alternatives delivered to us by contingent scientific practice (1989, chapter 6, §4); and even that it violates probabilistic coherence (1989, chapter 7, §4). The details of van Fraassen's criticisms are not important here, but their general tenor is worthy of note. Nancy Cartwright is not unfair in giving the following caricature of van Fraassen's challenge:

. . . what reason do we have for inferring from the fact that a bundle of principles save the phenomena to the fact that they are true? We need some reason, some good reason, though certainly not a conclusive reason. Many arguments wear their validity on their sleeve: 'I think. Therefore, I exist.' But not, '*P* explains *Q*. *Q* is true. Therefore *P* is true.' (Cartwright 1983, p. 89)

There are two possible premisses which might be used to give this argument its validity. The first would be that the world conspires to structure itself so as to reflect our explanatory relations. This is a strong metaphysical assumption; it is hard to see what could warrant it. (For a critique of such assumptions, see van Fraassen (1989), pp. 142–149.) The second, more palatable, premiss is that explanations of necessity reflect the structure of the world; it is the nature of explanation that *P* does not explain *Q* unless *P* is true.

It is hardly surprising, then, that the last decade has seen a move away from theories of explanation which are neutral on the question of scientific realism, as Hempel's largely was, to theories which issue from bigger philosophical pictures of the scientific enterprise. Realists and anti-realists have become divided over the issue of scientific explanation. (For a fuller discussion of this point, see Railton (1989).) For Wesley Salmon, a leading proponent of scientific realism, to give an explanation of an event is to point to its causes and to show how it fits into a causal nexus. (See, e.g., Salmon (1984).) So, according to Salmon, explanations do reflect the causal structure of the world, giving the argument '*P* explains *Q*; *Q*; therefore *P*' some force. Van Fraassen, by contrast, takes an explanation to be an answer to a why-question that satisfies certain constraints determined by the context in which the question was asked. (Van Fraassen 1980, Chapter 5.)

Nevertheless, it is not obvious that constructive empiricism could not incorporate a causal theory of explanation into its picture of science. Consider the following elementary claim that is made within the (presumably accepted) atomic theory:

Water consists of molecules; each water molecule is made up of one oxygen atom and two hydrogen atoms.

Adherents of earlier brands of anti-realism, such as positivism or instrumentalism, would have to reinterpret this claim, say as one about observable consequences, or as a tool devoid of truth value. Van Fraassen interprets it literally, but is anti-realist in his epistemic attitude toward it: he withholds his belief from it. In this way, van Fraassen avoids many of the problems associated with earlier forms of anti-realism.

Consider a related claim that might also be made within the atomic theory:

The molecular structure of water and of salt explains the dissolution of salt in water.

A theory of explanation should tell us how to interpret this claim. While it is not clear what would constitute a *literal* interpretation of this explanatory claim, a natural reading takes it to be a *causal* claim which entails the existence of atoms and molecules. Is the anti-realist thereby debarred from giving this reading to the claim? No – he can treat this explanatory claim exactly as he did the earlier descriptive claim, interpreting it realistically, but adopting an anti-realist epistemic attitude toward the claim. There is no *prima facie* reason why the anti-realist cannot adopt any interpretation of explanatory claims that is available to the realist.

This position will not be open to the anti-realist, however, if it can be shown that on certain interpretations of explanation – say causal accounts – explanations secure belief in unobservable entities. This is precisely what Cartwright aims to show.

II

Cartwright argues for at least three different but related claims (although she does not always distinguish them):

- (i) 'In causal explanations truth is essential to explanatory success.' (1983, p. 10.)
- (ii) '... inference from effect to cause is legitimate.' (p. 89.)
- (iii) 'To the extent that we find the causal explanation acceptable, we must believe in the causes described.' (p. 5.)

Claim (i) is ontological, whereas claims (ii) and (iii) are epistemological. Claim (i) says roughly that '*P*' does not causally explain '*Q*' unless '*P*' is true. Considerations raised in Section I above suggest that this claim is not in itself incompatible with an anti-realism along the lines of van Fraassen's constructive empiricism. Claims (ii) and (iii), by contrast, present direct challenges to anti-realism – (ii) by claiming the legitimacy of inference to the best causal explanation, and (iii) by providing an instantiation of the realist's thesis that acceptance involves belief – so we must take care to distinguish the two claims and examine each carefully. In particular, we must question whether either claim follows from claim (i).

Cartwright argues for these claims in two different ways, supporting them with both positive and negative arguments. While the former will be the focus of our attention, a digression on the latter will be useful. These arguments are negative in the sense that they attempt to block a certain line of argument against claims (i), (ii), and (iii); they do not directly support these claims.

It is common in the philosophy of science to distinguish between the observational and theoretical components of a theory, but this is not the distinction which Cartwright, in her (1983), takes to be primary. Rather, she distinguishes between fundamental and phenomenological laws. Phenomenological laws accurately describe 'low-level' regularities, and are typically discovered experimentally, rather than derived from theoretical considerations. Phenomenological laws need not describe observable regularities: 'the half-life of Carbon-14 is 5730 years' is a perfectly acceptable phenomenological law. Fundamental laws are high-level statements within a theory, typically differential equations; Schrödinger's equation and Einstein's field equations are paradigms.

Cartwright also distinguishes between two kinds of explanation: causal and theoretical. Causal explanations cite events that lead up to the event to be explained. Causal explanations will frequently involve unobservable entities and their properties. Such explanations are not typically committed to any particular theoretical treatment of the unobservable entities involved but only to their 'low-level' properties. Roughly speaking, causal explanations are committed to certain phenomenological laws, but not to any fundamental laws. Thus, a causal explanation of the results of an electroplating experiment will cite the charge of an electron, but not a classical or quantum mechanical theory of the electron.

Cartwright presents a novel account of theoretical explanation which she calls the 'simulacrum' account. On this account, explanation proceeds in roughly four stages. Suppose we have a physical system which exhibits a certain feature that can be described by a (probably messy) phenomenological law, and that we want to explain this feature. First, we choose a model (a simulacrum) of the physical situation. Next, we determine the mathematical properties of the system as represented in the model. These two steps are governed by the bridge principles of the theory we will use to provide the explanation; these principles will tell us which models are appropriate for which physical situations, and what their mathematical properties are. The third step is to plug the

mathematical properties of the system into one of the fundamental equations of the theory. Finally, one 'derives' an analogue of the phenomenological law governing the system in question. The scare quotes signify that the derivation is rarely mathematically rigorous, typically involving approximations, and corrections supplied by auxiliary phenomenological laws. The result of the derivation is an analogue in the sense that it describes the model in which the derivation was carried out, and not necessarily the physical system itself.

Suppose, for example, that we wish to explain the energy spectrum of the hydrogen atom. Our first step might be to model the hydrogen atom as an electron with mass m_e , electric charge $-e$, and spin S following a circular orbit at a distance r from a proton with mass m_p and electric charge $+e$. We will assume that there is no external magnetic field acting upon the atom, and we will ignore the spin of the proton. The second step is to determine the Hamiltonian of this system, since the Hamiltonian is the operator corresponding to the energy of a system (this is one of the bridge principles of quantum mechanics). The Hamiltonian for such a system would contain several components. Let \mathbf{p} be momentum of the electron in the frame of reference which takes the centre of mass of the atom to be fixed (thus $-\mathbf{p}$ is the momentum of the proton). Then the kinetic energy of the electron is

$$\mathbf{p}^2/2m_e - \mathbf{p}^4/8m_e^3c^2 + \mathbf{p}^6/32m_e^5c^4 - \dots$$

and similarly for the proton, where all but the first term represent relativistic corrections. It will suffice to consider only the first two terms of the expression for the electron, and the first term for the proton.³ Another term in the Hamiltonian will be the electric potential $-e^2/r$. Finally, there will be a term representing the interaction between the magnetic moment of the electron due to its spin, and the magnetic field 'seen' by the electron due to the relative motion of the proton around the electron.⁴

In order to determine the energy eigenvalues, we substitute this Hamiltonian into the time-independent Schrödinger equation: $\mathbf{H}u_E = Eu_E$, where u_E is an energy eigenfunction with eigenvalue E . This is the third step. Finally, the equation must be solved. It cannot be solved exactly for the Hamiltonian we have constructed, but an exact solution is available for the slightly simpler Hamiltonian $\mathbf{H}_0 = \mathbf{p}^2/2\mu - e^2/r$, where $1/\mu = 1/m_e + 1/m_p$.⁵ This is the Hamiltonian \mathbf{H} without the relativistic correction term for the kinetic energy of the electron, and

without the term for spin-orbit interaction.⁶ Using the solution for H_0 , we can get an approximate solution to the equation $Hu_E = Eu_E$ using time-independent perturbation theory. This will give approximate energy eigenvalues for the system described in our original model which will closely match the energy levels of a hydrogen atom that is not exposed to a strong external magnetic field.

The power of physical theories, according to Cartwright, comes from their ability to explain a diverse body of phenomena with as few fundamental laws, and as few bridge principles as possible. This demand for breadth on the part of the explanatory principles runs counter to the demand for their accuracy: if one wanted true explanatory principles one would employ a myriad of bridge principles and phenomenological laws, rather than seek unity through minimizing the number of principles used. Fundamental laws gain their generality by describing *models* which may be applied to many different real systems, of which the laws are literally false. Two important features of theoretical explanations are, therefore:

- (1) They involve models, which are abstractions of real phenomena, and the laws appearing in an explanation are true of the objects in the model, not of the objects in the real world.
- (2) The explaining principles of a good theory are few in number and broad in scope.

Both of these features argue *against* the truth of the principles figuring in theoretical explanations, including the fundamental laws. A third important feature of theoretical explanations is that science tolerates alternative theoretical explanations of the same phenomenon. Thus:

- (3) Alternative theoretical explanations of the same phenomenon complement, rather than compete with, one another.

Cartwright illustrates this point with the multiple theoretical treatments of quantum damping phenomena (1983, pp. 78–81). This feature of theoretical explanations also argues against the truth of the explaining principles: if truth were a pre-requisite to explanation, explanatory diversity should not be tolerated.

These arguments do not apply to causal explanations, according to Cartwright. Causal explanations describe actual states of affairs leading up to particular events, not abstract models. These explanations are not meant to have unifying power, so they can be tailor-made to fit

particular explananda. Thus, in the case of causal explanation, there is no desideratum such as unifying power which is at odds with the desideratum of truth. Moreover, in the case of causal explanation, there is no tolerance for a diversity of explanations:

We can infer to the truth of an explanation only if there are no alternatives that account in an equally satisfactory way for the phenomena. In physics nowadays, I shall argue, an acceptable causal story is supposed to satisfy this requirement. But exactly the opposite is the case with the specific equations and models that make up our theoretical explanations. There is redundancy of theoretical treatment, but not of causal account. (Cartwright, 1983, p. 76)

So, despite the multiplicity of theoretical explanations of quantum damping, it is generally agreed that quantum damping effects are *caused* by the emission and absorption of photons.

Cartwright's arguments are challenging, and deserve attention. It is not clear that causal explanations, as a rule, lack the three features that block the move from explanatory power to truth in the case of theoretical explanations. With regard to the first feature, Cartwright's own recent work on capacities (1989a, 1989b) suggests that some causal explanations do involve a level of abstraction. As to the second feature, genetics, for example, is a science that seems to be in the business of describing a small number of causal mechanisms that have great explanatory breadth. Philip Kitcher (1989) has even proposed a reduction of the concept of causation to explanatory unification. Third, Paul Humphreys (1989a, see also 1989b) argues persuasively that legitimate causal explanations need not be complete. If this is correct, then there will be many causal explanations (citing different combinations of causal factors) of any single event. Rather than pursue these objections further, I will simply reiterate the point that the arguments of Cartwright considered so far are purely negative: they aim to show that the arguments adduced against inference to the best *theoretical* explanation do not undermine inference to the best causal explanation; even if successful they do not show that inference to the best causal explanation is legitimate.

What, then, is Cartwright's positive argument for claims (i), (ii), and (iii)? Her argument for claim (i) is straightforward: in causal explanation, unlike theoretical explanation, truth is not an optional extra; but a necessary ingredient. Her argument for claims (ii) and (iii) rely on an appeal to our intuitions. She invites the reader to imagine himself in the epistemic situation described in the antecedent of (iii), accepting

a causal explanation, and asks if he does not then believe in the purported cause:

My newly planted lemon tree is sick, the leaves yellowing and dropping off. I finally explain this by saying that water has accumulated in the base of the planter: the water is the cause of the disease. I drill a hole in the base of the oak barrel where the lemon tree lives, and foul water flows out. That was the cause. Before I had drilled the hole, I could still give the explanation and to give that explanation was to present the supposed cause, the water. There must be such water for the explanation to be correct . . .

God tells you that Schrödinger's equation provides a completely satisfactory derivation of the phenomenological law of radioactive decay. You have no doubt that the derivation is correct. But you still have no reason to believe in Schrödinger's equation. On the other hand, if God tells you that the rotting of the root is the cause of the yellowing of the leaves, or that the ionization produced by the negative charge explains the track in the cloud chamber, then you do have reason, conclusive reason, to believe that there is water in the tub and that there is an electron in the chamber. (1983, 91-3)

I do not deride these arguments by calling them appeals to intuition. The thought experiments described provide compelling arguments in support of Cartwright's epistemological claims, and must be treated with respect.

III

This section will examine claim (ii): ' . . . inference from effect to cause is legitimate.' What is meant by the claim that this inference is 'legitimate'? Plausibly, it could mean that the inference from '*Q*' and '*P*' causally explains *Q*' to '*P*' is valid. This would make claim (ii) equivalent to claim (i), which says that truth is necessary for causal explanation. But the validity of inference to the best causal explanation is not enough for the realist to infer to the existence of unobservable entities; the inference must be *sound*.⁷ I will argue that validity can only be purchased at the expense of soundness; if causal explanation presupposes truth, then the anti-realist's agnosticism about the truth of any sentence '*P*' which purports to describe unobservable entities will commit him naturally to an agnostic attitude toward claims of the form '*P* causally explains *Q*.'

An analogy will help to set the stage. The realist and the anti-realist disagree over whether '*Q*; *P* explains *Q*; Therefore *P*' is a scheme of inference that allows us to infer the existence of unobservables. The debate would seem to hinge on whether 'explain' is a success verb: can '*P*' explain '*Q*' if '*P*' is not true? A familiar success verb is 'know'.

Horace cannot know that *P* unless '*P*' is true; so 'Horace knows that *P*, therefore *P*' is a valid inference. By contrast, 'believe' is not a success verb; Horace can believe that *P*, while '*P*' is false. Realists have typically taken 'explain' to function like 'know', and anti-realists have taken it to function like 'believe'. Cartwright, we have seen, argues that theoretical explanation functions like belief, and causal explanation like knowledge. But is any concession made by the anti-realist who agrees that 'explain' is a success verb? Not if the belief/knowledge analogy is any guide. There is, after all, no such thing as 'inference from knowledge'; the inference pattern 'Horace knows that *P*, therefore *P*', although valid, is of no use to anyone, not even Horace. By examining Horace, the best anyone (including Horace) can determine is that Horace *believes* that *P*. In order to collect evidence that Horace *knows* that *P* and doesn't just believe it, we must first collect evidence that '*P*' is true. Once we have done that, what need have we to infer from Horace's knowledge?

Recall Cartwright's argument:

My newly planted lemon tree is sick, the leaves yellowing and dropping off. I finally explain this by saying that water has accumulated in the base of the planter: the water is the cause of the disease . . . There must be such water for the explanation to be correct. (1983, p. 91)

Doesn't this just show that we are not in a position to accept the explanation of the ailing citrus tree *until* we have good reason to believe that there really is water in the base of the planter? If so, the fact that the water in the planter *explains* the tree's poor health cannot provide our epistemic warrant for believing the water to be there, for we must have such warrant already if we are to accept the proposed explanation.

The matter will be clarified if we divide the causal explanation into two parts: the causal story, or description of the cause, and the events and entities which make the causal story true. Cartwright focusses on the second part – without the postulated entities there would be no explanation. But what scientists have at their disposal when forced to make epistemic commitments is the first part, the causal story. Imagine a physicist or chemist working shortly after the turn of the century who is interested in the phenomenon of Brownian motion. It was known from the theoretical work of Smoluchowski and Einstein, and from the experimental work of Perrin, that the atomic theory presented a potential explanation of the peculiar motions of small particles suspended in

a liquid (typically these particles themselves were not 'observable' but only seen through high-powered microscopes – we ignore this difficulty): the particles were being bombarded by molecules, and thus were constantly gaining linear momentum in random directions. Was such a working scientist in a position to believe this explanation? There are two possible cases. First, she might already have good reason to believe in the molecules which figure in the causal explanation. If this were the case, she would have no difficulty in accepting the proffered explanation. But this could not be an instance of inference to the best explanation of the sort that lends any comfort to the realist; *ex hypothesi* our scientist already had good reason to believe in molecules, so the causal explanation in question played no role in her coming to believe that molecules are real. Suppose, then, that she previously did not have strong reason to believe in molecules; by what criterion could she judge the proffered causal explanation? One thing is clear: she has to base her acceptance on the strength of the causal story alone – she is, as yet, uncertain about the entities that would make the story true. There are many features of a causal story which might lead us to accept it as a good one: it might be simple, self-consistent, and applicable to a broad class of known phenomena; it might resemble causal explanations of related or analogous phenomena which we already accept, generate accurate predictions, and so on. But a challenge remains unanswered: why do these features of a causal story warrant the belief that the story is true? Put another way: how can any features of the causal story alone allow us to infer that it is a genuine causal *explanation*? The original challenge to inference to the best explanation is *not* met in the case of causal explanations, but only displaced. It was asked: "What is it about explanation that warrants the inference from '*Q*' and '*P* explains *Q*' to '*P*'?" In the case of causal explanation, Cartwright argued, there is an answer. But the question re-surfaces: "What is it about a (good) causal story that warrants the inference from 'there is a (good) story according to which *P* causes *Q*' to '*P* causes *Q*'?" Since the working scientist has access only to features of the causal story, such an inference must be warranted if she is ever to discharge the premiss '*P* causally explains *Q*'. The advocate of inference to the best explanation is faced with a dilemma: in order to accept the premiss that *P* causally explains *Q*, the scientist must either (i) have independent grounds for believing that the cause exists, in which case the purported causal explanation is idle with respect to any inference to the existence of the cause; or (ii) she

must make an inference based on features of the causal story alone, in which case an argument is required if we are to accept the inference as valid.⁸

In fairness to Cartwright, she is sensitive to some of the objections I have been raising. Cartwright agrees that it is not legitimate to infer to the truth of a causal story based only on features of the story, such as simplicity, internal consistency, and so on. Rather, we should only infer to the reality of the cause once we have determined that the effect could not have been brought about in any other way:

We must have reason to think that this cause and no other, is the only practical possibility, and it should take a good deal of critical experience to convince us of this.

We make our best causal inferences in very special situations . . . where the likelihood of other causes is ruled out . . . Seldom outside of the controlled conditions of an experiment are we in a situation where a cause can legitimately be inferred. (Cartwright, 1983, p. 6)

We rule out alternative causal candidates by producing effects in controlled conditions, where we know alternative causes cannot be at work. But this answer does not escape the dilemma, it merely chooses a horn to sit on; it buys the validity of inference from causal story to cause at the cost of postulating that there are independent means of finding out about the cause. It would beg the question against the realist to deny that there are experimental means of discovering what unobservable causes are at work in producing a certain effect. If such means exist, then perhaps inference to the best causal explanation can play some role in the realist's vision of science; but inference to the best causal explanation is not the realist's point of entry into the unobservable realm on Cartwright's picture, for its application presupposes means for finding out about the unobservable:

. . . unlike theoretical accounts, which can be justified only by inference to the best explanation, causal accounts have an independent test of their truth: we can perform controlled experiments to find out if our causal stories are right or wrong. (Cartwright, 1983, p. 82.)

If we can, indeed must, perform such experiments, aren't we inferring from the results of the experiment, rather than from the causal explanation?

Cartwright is also aware of the strategy I have recommended to the anti-realist: avoid inferring from '*Q*' and '*P* explains *Q*' to '*P*' by refusing to embrace the second premiss. Cartwright, however, seems to view

this refusal as an extra sceptical move on the anti-realist's part, rather than one to which he is already committed. Concerning van Fraassen's rejection of such causal explanations, Cartwright writes:

Van Fraassen does not believe in causes. He takes the whole causal rubric to be a fiction – that is irrelevant here. Someone who does not believe in causes will not give causal explanations. One may have doubts about some particular causal claims, or, like van Fraassen, about the whole enterprise of giving causal explanations. These doubts bear only on how satisfactory you should count a causal explanation. They do not bear on what kind of inferences you can make once you have accepted that explanation. (Cartwright 1983, p. 93.)

Of course, an anti-realist can be a sceptic about causation in general, but he needn't be. His scepticism toward certain causal explanations is a natural consequence of his agnosticism toward theoretical entities. If 'P causally explains Q' cannot be true unless 'P' is true, the anti-realist may simply exercise his right to apply *modus tollens* rather than *modus ponens*.

I wish to end this section with a brief discussion of the famous experimental work of Jean Perrin. Cartwright cites Perrin as a practitioner of inference to the best causal explanation *par excellence*. In Cartwright's reconstruction, Perrin began with a causal hypothesis – that Brownian motion in tiny particles suspended in a liquid is the result of constant bombardment by molecules. He then constructed an experiment that would be highly sensitive to the nature of the postulated causes. He observed the vertical distribution of colloid particles at equilibrium. Based on his detailed observations of particle densities at different heights, he was able to employ his causal model to determine Avogadro's number: the number of molecules in a mole of a substance. Thus, Perrin inferred to the best causal explanation of the observed vertical distribution: one mole of a substance is composed of roughly 6×10^{23} molecules. According to Cartwright, however, such an inference to the most likely cause is only warranted if alternative explanations can be ruled out. Perhaps the vertical distribution was a function of the experimental set-up, or of the peculiarities of the liquid in which the particles were suspended. In order to allay such fears, Perrin reports twelve other experiments designed to determine Avogadro's number (several of them performed by Perrin himself). Each experiment involved different sorts of apparatus, different materials, and different physical and chemical effects. All yielded results ranging from 6 to 7.5×10^{23} molecules per mole (Perrin 1923, pp 206–7). If the cause of

Perrin's observations had been in the experimental set-up, then analogous but independent causes would have to be responsible for the other results. Such a conspiracy to produce comparable results in such diverse experiments is so improbable that it can be dismissed as a serious possibility. Thus, Perrin's inference to the most likely cause was indeed legitimate.

The case of Perrin has been much discussed in the literature,⁹ and I wish to add as little as possible here. I simply offer two alternative reconstructions of Perrin's argument that strike me as more compelling than Cartwright's. First, in Bayesian confirmation theory, a hypothesis receives confirmation when its probability is increased as a result of conditionalization on evidence. By Bayes' theorem, $P(H/E) = [P(H) \cdot P(E/H)] \div P(E)$, so whenever $P(E/H) \div P(E) > 1$, E provides confirmation of H. The atomic theory predicts that if the thirteen experiments described by Perrin are performed, and the results used to calculate Avogadro's number, there is a high probability that all the results will be within a certain limited range, say that there is a number x so that all results fall within $x \pm (x/8)$. Moreover, this result is much less likely if the atomic theory is false, indeed there would be no reason to expect calculations based on experiments on Brownian motion, X-ray diffraction, α -decay, and the colour of the sky to be correlated at all. Thus $P(H/E) > P(H)$. Since the predictions of the atomic theory were borne out, it was confirmed by the experiments. A second interpretation is provided by Glymour's bootstrapping theory of confirmation (Glymour 1980). Roughly put, evidence E confirms a hypothesis H with respect to background theory T if E and T together allow a 'computation' of H, and if different evidence would have allowed a computation of $\neg H$. Here, the background theory is the atomic theory. H the hypothesis that one mole of a substance contains a constant number of molecules, and E the results of at least two of the experiments referred to by Perrin. It is easy to check that T, H, and E (roughly) fit the bootstrapping schema in this example. In particular, had the two experiments yielded significantly different values for Avogadro's number, $\neg H$ could have been deduced. In neither of these reconstructions is any appeal made to causal explanation.

The work of Perrin is impressive, and his argument for the existence of atoms persuasive. Perhaps commitment to belief in unobservables can be secured by confirmation fitting either the Bayesian or the Bootstrapping schema – that is another debate. My claim in this section

pertains only to one purported pattern of confirmation: inference to the best causal explanation. Unfortunately, in any attempt to apply the pattern, the inference must be either invalid or superfluous. Cartwright's example of Perrin's argument for the existence of atoms is a case where inference to the best explanation is superfluous: the argument can be reconstructed without any allusion to the causal explanations provided by Perrin's atoms.

IV

We turn now to Cartwright's claim (iii): '[t]o the extent that we find the causal explanation acceptable, we must believe in the causes described'. Before examining Cartwright's argument for (iii), I wish to dismiss the claim that (iii) follows easily from (i) – '[i]n causal explanation truth is essential to explanatory success' – by recasting the central argument of the previous section. The seemingly easy step from (i) to (iii) rests on an ambiguity in the phrase 'explanatory success'. If success is read semantically – in the sense in which we talk of 'know' being a 'success' verb – then Cartwright is entitled to claim (i): where there is no real cause, there is no causal explanation. But we are tempted to read 'success' pragmatically; Nineteenth Century explanations of optical phenomena which cited the propagation of waves through the aether were successful, as were explanations of thermodynamic phenomena in terms of the flow of calorific fluid. In this sense, success does not require truth, but only usefulness, relevance, psychological satisfaction, and so on. What is the relationship between acceptance and success? If we are to infer from (i) to (iii), something like the following must be true: 'acceptance of a causal explanation entails the belief that it is successful'. Is this premiss plausible? Yes, if we give the second, pragmatic reading to 'success', but this is the reading which renders (i) false. It is not at all clear that the premiss linking acceptance and success is true on the first reading of 'success'; in any event, this reading of the premiss comes close to simply denying van Fraassen's position, so it cannot be accepted without careful argument. The challenge now facing the realist can be framed by once again distinguishing between a causal story and the truth-makers in the world which make the story a genuine explanation. What we accept is not the explanation *per se*, but the causal story. Cartwright herself employs this sort of language:

If we accept Descartes' causal story as adequate, we must count his claims about hooked atoms and vortices true. (Cartwright 1983, p. 75)

In physics nowadays, I shall argue, an acceptable causal story is supposed to satisfy this requirement. (Ibid., p. 76)

The challenge: granted that a causal story is only a genuine explanation if it is true, whence does it follow that to accept a causal story is to believe that it is a genuine causal explanation?

Cartwright argued for claim (iii) by means of a *Gedankenexperiment* involving Divine diagnosis of her lemon tree's ailment. Indeed, if God tells us that water in the tub is the explanation of the lemon tree's condition, we had better accept this explanation. But must we *believe* it? We might believe it on the Cartesian principle that God could not be a deceiver. Locke, however, held that God hid from us the actual micro-structure of objects, instead 'super-adding' secondary qualities such as colour and flavour onto the effects the objects would normally have upon our sense organs. (See Locke (1975), II.viii, II.xxxi, and IV.iii *inter alia*.) Locke's Deity did this not out of a malicious desire to deceive, but out of concern for His creatures, flagging poisonous berries with distinctive colours and odours, rather than leaving us to calculate the effects of ingesting the berries from their primary qualities, such as the size, shape, and solidity of their constituent corpuscles. While God might 'tell us' that the berries are green by adding this secondary quality to our perception of them, we should not believe that there really is a quality 'green' inhering in the berries. God's superaddition is to aid us in our struggle for survival, which must, alas, take precedence over our quest for knowledge of the world. This discussion is of course fanciful – the debate over scientific realism does not turn on a sticky point of theology. A more fruitful discussion ensues if we turn from causal explanations delivered from on high by the Almighty to those delivered in textbooks by highly respected, albeit mortal, scientists.

Consider an explanation of the uncertainty relations proffered in many introductory and intermediate level textbooks on quantum mechanics. (I will follow the treatment of Gasirowicz (1974).) Imagine the following, rather idealized set-up to measure the position of an electron: A stream of electrons moves from left to right under a screen (see Figure 1). We fire a photon of light with wavelength λ from right to left to collide with an electron. The photon will be deflected and (if we are lucky) pass through a lens to be registered on a screen. The

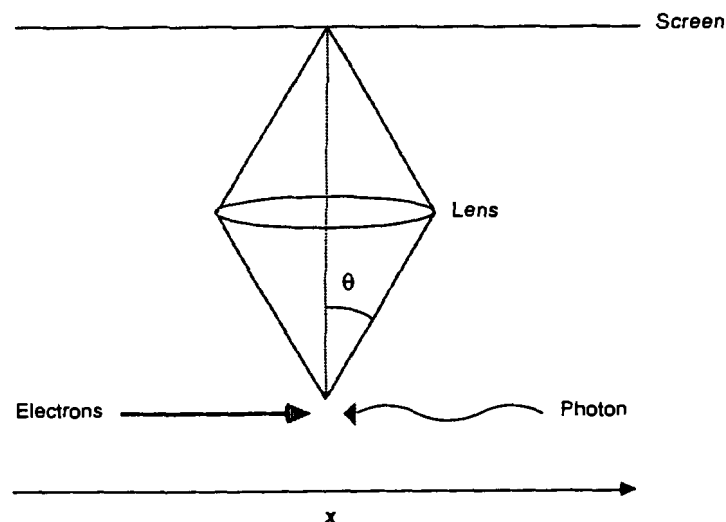


Fig. 1. A set-up to measure the position and momentum of an electron. A photon is fired from right to left to collide with an electron. If the photon recoils and passes through the lens, it will register on a screen, and the location of the collision can be roughly determined. In order to increase the accuracy of the measurement, the frequency of the photon and the angle subtended by the lens, θ , must be increased. These changes decrease the accuracy of the measurement of the electron's momentum.

precision with which the position (we are here concerned only with the one-dimensional position along the x -axis) of the electron can be measured can be determined by classical optics: $\Delta x \approx \lambda / \sin \theta$, where 2θ is the angle from the point of collision subtended by the lens. In order to make the position measurement more precise, we must use light of shorter wavelength. Since $\lambda\nu = c$, where ν is the frequency of the light and c is the velocity of light, a constant, decreasing the wavelength of the photon involves increasing its frequency. The momentum of the photon is given by $h\nu/c$, where h is Planck's constant, so momentum is inversely proportional to wavelength. We can measure the momentum of the electron by measuring the change in momentum of the photon after it scatters off the electron. But due to the distorting effect of the lens, we can only know the direction of scattering of the photon to within an angle of $\pm\theta$. If the photon scatters with an angle of ϕ to the vertical, the x -component of its momentum will be proportional to $\sin \phi$, but it will also be proportional to its absolute momentum $h\nu/c$.

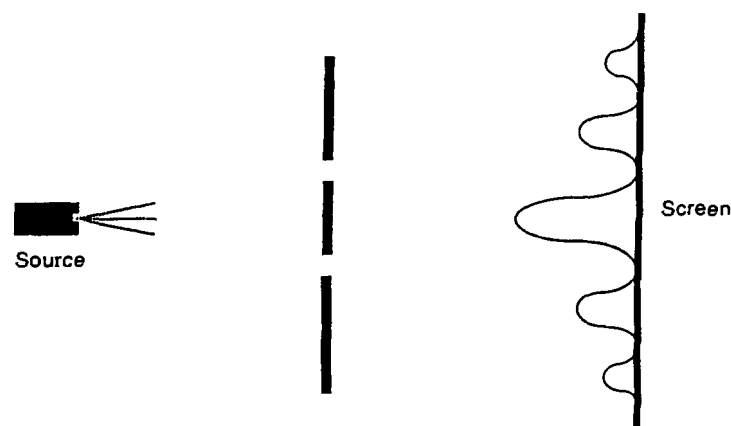


Fig. 2. The two-slit experiment. Electrons are fired slowly at a screen with two apertures. The electrons that pass through the apertures are registered on a second screen. The sine curve represents the density of electron hits in each region of the screen. The pattern suggests the interference of waves propagating through the apertures.

So $\Delta p_x \approx 2(h\nu/c) \sin \theta$. The position measurement is accurate when we use light with short wavelength and high frequency, whereas the momentum measurement is accurate when we use light with low frequency and long wavelength. More formally, we have

$$\Delta p_x \Delta x \approx 2(h\nu/c) \sin \theta (\lambda / \sin \theta) \approx 2h,$$

agreeing with the uncertainty relation $\Delta p_x \Delta x \geq h/2\pi$ predicted by quantum theory.

A similar explanation is frequently applied to the two-slit experiment, which is designed to exhibit the dual wave and particle nature of light as well as of bodies. Imagine that a source fires electrons at a screen with two slits in it, with those electrons passing through the slits being registered on a second screen (Figures 2 and 3). (In practice this sort of experiment would be too crude for electrons; instead the electrons are bounced off the surfaces of crystals composed of layers of molecules aligned at an oblique angle to the trajectory of the electrons. One then looks for interference patterns between those photons deflected by different molecular layers.) What the experiment shows is that the distribution of hits on the second screen when both slits were open exhibits an interference pattern (Figure 2). In particular, the pattern is not what one would get by superimposing the distributions resulting

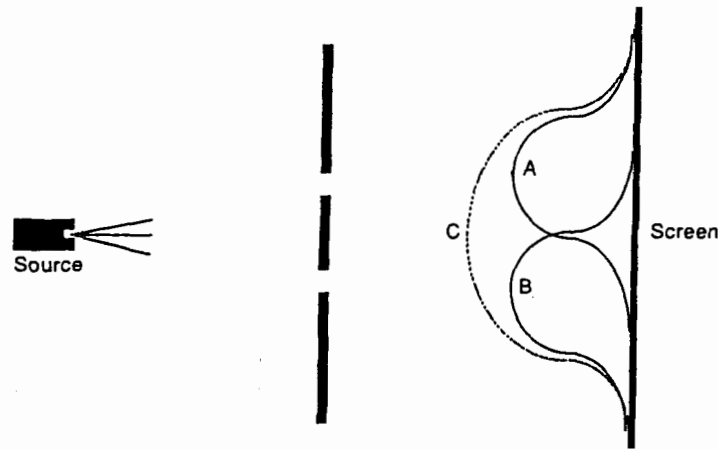


Fig. 3. The two-slit experiment with one aperture covered. Curve A represents the pattern of electron hits with the bottom aperture closed; curve B represents the pattern of hits with the top aperture closed. Curve C is the result of superimposing curves A and B. Note the difference between curve C and the pattern of hits shown in Fig. 2, where both slits are open. If both slits are open, but the position of each electron is measured as it passes through the first screen, the pattern of electron hits will resemble curve C rather than the interference pattern of Fig. 2.

from the two versions of the one-slit experiment (Figure 3). This interference pattern results even if the source emits electrons at an extremely slow rate, say one per second. Our classical intuitions tell us that each electron must go through one slit or the other, and if the electrons are being emitted too slowly for subsequent electrons to be interfering with one another, we would expect the result of the two-slit experiment to simply be the superposition of the results of the two one-slit experiments. This suggests the following paradox: Suppose during the running of a two-slit experiment, we determine which slit each electron passes through. Then we examine the distribution of hits on the screen for all the electrons that passed through the top slit, and all those which passed through the bottom. Each should give rise to a single slit distribution, but together they should give rise to the two-slit distribution. How is this possible? Suppose that we try to determine which slit a given electron passes through by bouncing a photon off of it as it passes through. The wavelength of the photon must be small enough that the position measurement on the electron after it has passed through the

first screen is accurate relative to the distance between the two slits. The collision of the photon with the electron will impart some momentum to the electron in the vertical direction, with the amount of momentum transmitted being proportional to the frequency of the photon. If the electron receives a sufficient kick, its vertical displacement upon reaching the screen will be of the same order of magnitude as the distance between the crests of the interference pattern, so the interference pattern will be destroyed. A simple calculation (omitted here) shows that if the photon used to measure the position of the electron has sufficiently small wavelength to determine which slit the electron passed through, then its frequency is sufficiently large to destroy the interference pattern. If you catch each electron going through one slit or the other, the two slit interference pattern will not result.

Here we have two causal stories that bear all the marks of accepted causal explanations. All except one: the proffered explanations contradict almost every interpretation of quantum mechanics, and as such, would not be believed as literally true by any but the most stubborn believer in hidden variable theories. The explanations given assume that the electron has definite position and momentum values, and that the uncertainty relations express our ignorance of these values, which results from our inability to measure position accurately while controlling the amount of momentum introduced into the system. According to quantum mechanical orthodoxy (and most quantum mechanical heresy), however, the electron does not have a simultaneously sharp position value and momentum value. The electron must be in a superposition of eigenstates of at least one of the observables; typically it will be in a superposition with respect to both. When a position measurement is made, the superposition collapses, and the electron's new state will be an eigenstate of the position operator (or at any rate, sharply peaked with respect to position if the measurement is imperfect). It is only as a result of the measurement that the electron obtains a definite position value at all. The uncertainty relation then follows from the incompatibility of position and momentum eigenstates – it can be derived from elementary wave mechanics or more abstract operator methods.¹⁰ The causal explanations described above depict quantum measurement as a neo-classical affair, involving the collision of photons and the microscopic particles to be measured: the only non-classical element involved is the quantization of photon energy expressed in the relation $E = h\nu$. Would that it were so simple! Alas.

this interpretation of quantum mechanics is untenable in light of the various arguments against hidden variable theories, especially the results due to Kochen and Specker (1967). These causal explanations, then, provide good candidates for counterexamples to the claim that acceptance of a causal explanation involves commitment to belief in the truth of the causal story. The causal explanations of the uncertainty relations would seem to be accepted, but not believed.

Two objections will be raised against this claim. I will address the less obvious one first. Grant, for now, the claim that members of the scientific community do accept the causal explanation of the uncertainty relations. They do not believe the entire causal story, the objection would run, but they do believe in the unobservable entities postulated in the story – the photons and the electrons – and it is the epistemic commitments of scientists with regard to the existence of theoretical entities which is here at issue. This objection hinges on a controversial point in the philosophy of language: to what extent are the *identities* of the entities figuring in the causal stories determined by the contents of those stories. Theories of reference may be roughly divided into two types, those which I will label 'user-friendly', and those which are not user-friendly. The user-friendly theories of reference include many varieties of 'direct reference' theories as well as their cousins, the causal theories of reference. Reference is user-friendly on these accounts, because the relationship between word and object is not mediated by our beliefs about the object, or the descriptions of the object which we happen to accept. In extreme versions of the user-friendly conception of reference, the entities in the world all but line up in front of us to receive names, much as all the birds and animals were brought before man to receive names in the second book of Genesis. According to such a theory of reference, a story could be wildly false of electrons, and yet succeed in being a false story *about* electrons. By the same token, one could, on such a view, reject the causal story about photon-electron interactions while maintaining belief in the characters figuring in the story. By contrast, according to the non-user-friendly theories of reference, such as descriptivist theories, whether or not a word successfully refers to an object depends crucially on the set of (believed) descriptions in which the word figures. According to such a theory, if our theories about (what we call) electrons were sufficiently off the mark, they would not simply be false theories of (actual) electrons, but theories about non-existent objects; our word 'electron' would not have

referred to anything at all. Similarly, one could not believe a causal story to be sufficiently false, and still believe in the entities figuring in the story. The entities figuring in our causal stories above might be renamed 'classical electrons'; few scientists today believe in the existence of these entities, although many believe in quantum mechanical electrons.¹¹ It is, therefore, not at all clear whether one could consistently believe that the causal stories told above are literally false, but that the entities figuring in them really exist. In any event, if reference is sufficiently user-friendly, realism about theoretical entities would be trivially true, but would cease to be an interesting thesis about science. Belief in electrons would be warranted, not because of the reliability of our scientific methods, but because we could be assured that something in the world was kind enough to step up and be named when the term 'electron' was first coined.

The second objection is that the causal explanations of the uncertainty relations are not really accepted as explanations, despite the fact that they so commonly appear in the textbooks used to train the next generation of physicists. Rather, it could be argued, the causal stories play a heuristic role. Here is a plausible story: The uncertainty relations are, to the novice, a rather surprising and unpalatable consequence of the formalism of quantum mechanics. Indeed, in the late 1920's and early 1930's, Einstein objected to the quantum theory on the basis of scepticism about the uncertainty relations. Bohr used thought experiments similar to the causal stories above to defend the quantum theory against Einstein's criticisms. (See Bohr (1949).) The causal stories show that the uncertainty relations are to be expected even without all the apparatus of the quantum theory: the uncertainty relations follow (in a rough way) from classical optics, together with the quantization of photon energy. The latter phenomenon was established by means of the photoelectric effect, and its discovery by Einstein in 1905 was one of the first steps on the road to the development of quantum mechanics. The causal stories show how easy it is to be carried further down the road once we have made the first step. The causal stories also serve as an intuition pump: they prepare us for the principle that measurement of a system typically must disturb the system. While the model of strict causal disturbance by the measuring apparatus on the system had to be given up in order to respond to the thought experiment of Einstein. Podolski, and Rosen (1935),¹² it continues to provide a useful heuristic to guide us as gently as possible from the classical to the quantum

world. It is in this capacity, it would be argued, that the causal story is accepted.

The second objection is more clearly on the mark than the first. The moral to be drawn from these examples is not simply that there are causal explanations which are accepted but not believed; to draw this moral would be to deny the objection its due. The moral to be drawn is that there are many causal stories which are employed by the scientific community for a variety of purposes – some of them are believed as literally true, some are not. All of these causal stories are accepted in the sense that members of the scientific community deem them appropriate for the role they play; even the just-so stories can be accepted as serving important heuristic roles. Thus there is nothing special about causal stories such that acceptance of them commits one to belief in their truth. The realist faces a new challenge: what is special about the *role of explanation* such that causal stories filling that role, and not some other, must be believed if accepted?

This challenge will not be easy to meet. In the second objection above, it was argued that the causal stories involving the uncertainty relations were not accepted as explanations *because* they served some other purpose within the physics community. This suggests that the way in which a particular causal story is deployed will determine its role; if so, pragmatic factors will, in part, determine whether or not a causal story fits the role of explanation. This allows the anti-realist to adapt a familiar line of argument (see, e.g., van Fraassen (1989), p. 192): there are different reasons for accepting causal stories; at least some of the reasons for accepting a causal story *as an explanation* are pragmatic rather than epistemic; only epistemic grounds for acceptance can be grounds for belief: therefore to accept a causal explanation is not the same thing as to believe it to be true. It is the preponderance of pragmatic virtues that leads us to accept the causal stories involving the uncertainty relations, and, according to Locke, God's 'story' about secondary qualities. Even in Cartwright's example involving Divine revelation, it is not obvious why the overwhelming reasons for accepting the explanation offered should also be reasons for believing it.

V

While I have defended anti-realism from a number of challenges, it is not my intent to advocate anti-realism. I claim simply that the theory

of explanation is the wrong place to look for an argument against anti-realism. In the case of causal explanation, the inference

$$\frac{Q}{P \text{ explains } Q} \\ P$$

has a claim to validity. But establishing this validity is not enough to counter the challenge originally posed to the realist. In Section III, I argued that the anti-realist's scepticism concerning the existence of unobservables would lead him to be sceptical of those explanation claims that would warrant inferences to the unobservable. The anti-realist need make no ad hoc moves, nor, indeed, any additional moves at all, in order to avoid having to discharge the second premiss. In Section IV I argued that reasons for accepting a causal explanation need not be reasons believing it; again, nothing more than the standard repertoire of anti-realist moves is needed.

Anti-realism is not only compatible with a causal theory of explanation, such a theory could be incorporated into the anti-realist's picture of science in an entirely natural way. Salmon (1989) has expressed the hope that the next decade will begin to see a *rapprochement* between the competing theories of scientific explanation. Divorcing issues in the theory of scientific explanation from prejudices imported from the orthogonal debate over scientific realism will be a step in the right direction.

NOTES

• I would like to thank Wes Salmon for reading earlier drafts of this paper. This paper is based upon work supported under a National Science Foundation Graduate Fellowship.

¹ The phrase is from Harman; see his (1965, 1968). The idea can be traced back at least as far as Peirce's 'abduction'.

² I will use '*P*' and '*Q*' to represent sentences as well as their nominalizations. Thus '*P*' might at different times stand for 'water freezes at zero degrees celsius', 'that water freezes at zero degrees celsius', or 'the freezing of water at zero degrees celsius'. This ambiguity should cause no harm, and will obviate the need for awkward constructions such as "that '*P*' is the case . . ."

³ A quick calculation can show that this approximation will yield correct energy levels to within an order of less than 10^{-8} .

⁴ This term will be $(e^2/2m_e^2c^2r^3)(\mathbf{L} \cdot \mathbf{S})$, where \mathbf{L} is the orbital angular momentum operator. (The factor of two in the denominator is the result of the Thomas precession effect.)

⁵ Converting the equation $\mathbf{H}_0 u_E = E u_E$ to polar co-ordinates we get

$$(-\hbar^2/2\mu)[(1/r^2)(r\partial/\partial r)(r\partial/\partial r) + (1/r)(\partial/\partial r) - (1/\hbar^2 r^2)L^2]u_E(r) - (e^2/r)u_E(r) = E u_E(r)$$

where \mathbf{r} is the position vector of the electron relative to the proton, and \hbar is $h/(2\pi)$, h being Planck's constant. By separation of variables, we can write $u_E(\mathbf{r}) = R_{Elm}(r)Y_{lm}(\theta, \phi)$ where r , θ , and ϕ are the three dimensional polar co-ordinates, and $Y_{lm}(\theta, \phi)$ is a simultaneous eigenfunction of the L^2 and L_z (z -component of L) operators, having eigenvalues $\hbar^2 l(l+1)$ and $\hbar m$, respectively. After solving for Y we have

$$(*) \quad (-\hbar^2/2\mu)[(1/r^2)(r\partial/\partial r)^2 + (1/r)(\partial/\partial r) - l(l+1)/r^2]R_{Elm}(r) - (e^2/r)R_{Elm}(r) = E R_{Elm}(r).$$

Solving this equation with a change of variables yields a recursive relation satisfied by a set of Laguerre polynomials. This enables us to solve exactly for R_{Elm} and thus for E .

⁶ These two terms provide corrections on the order of 10^{-4} .

⁷ This objection is hinted at in Mayo (1986).

⁸ A similar point is made in Bogen and Woodward (1988, §VIII). They argue that the principle of inference to the best explanation may be interpreted in two ways. According to the weaker reading, a piece of evidence cannot ultimately support a hypothesis unless that hypothesis figures in a potential explanation of the evidence. On this reading, explanation is necessary but not sufficient for inference. According to the stronger interpretation, the explanatory power of a hypothesis is itself a reason for believing the hypothesis. Bogen and Woodward do not object to the weaker version of inference to the best explanation, but argue that many legitimate inferences in science do not in any way conform to a pattern of inference to the best explanation along the lines of the stronger reading.

⁹ A treatment that is relevant to my concerns here is given in Mayo (1986). Mayo argues that the causal inference described by Cartwright could only be made once it was established that the general theoretical treatment of Einstein and Smoluchowski was applicable. This required a previous *theoretical* inference of the sort which Cartwright eschews.

¹⁰ See, e.g., Gasiorowicz (1974), chapter 2 and appendix B.

¹¹ It was noted in Section II that, according to Cartwright, one can typically give a causal explanation involving electrons without being committed to a particular theoretical treatment of electrons. Here, however, we have a case where the theoretical treatment finds its way down to the level of properties figuring in the causal explanation. The causal details of measurement interactions depend crucially upon whether electrons have simultaneously sharp position and momentum values, or whether they are typically in states of superposition.

¹² In the EPR experiment, a measurement performed on one particle influences the result of measurement on a distant particle. Bohr responds: 'It is true that in the measurements under consideration any direct mechanical interaction of the system with the measuring agencies is excluded, but a closer examination reveals that the procedure of measurements has an essential influence on the conditions on which the very definition of the physical quantities rests' (Bohr 1935a). Bohr uses very similar language in his

(1935b). For a discussion of the evolution of Bohr's conception of disturbance in response to Einstein's challenges, see Fine (1981).

REFERENCES

- Barker, P. and C. G. Shugart (eds.): 1981, *After Einstein*, Memphis State University Press, Memphis.
- Bogen, J. and J. Woodward: 1988, 'Saving the Phenomena', *Philosophical Review* 97, 303-352.
- Bohr, N.: 1935a, 'Quantum Mechanics and Physical Reality', Letter to *Nature* 135, 65. Reprinted in Wheeler and Zurek (1983).
- Bohr, N.: 1935b, 'Can Quantum Mechanical Description of Physical Reality Be Considered Complete?', *Physical Review* 48, 696-702. Reprinted in Wheeler and Zurek (1983).
- Bohr, N.: 1949, 'Discussion with Einstein on Epistemological Problems in Atomic Physics', in Schilpp (1949), pp. 200-41. Reprinted in Wheeler and Zurek (1983).
- Cartwright, N.: 1983, *How the Laws of Physics Lie*, Oxford University Press, Oxford.
- Cartwright, N.: 1989a, *Nature's Capacities and Their Measurement*, Oxford University Press, Oxford.
- Cartwright, N.: 1989b, 'Capacities and Abstractions', in Kitcher and Salmon (1989), pp. 349-356.
- Churchland, P. and C. Hooker (eds.): 1985, *Images of Science*, University of Chicago Press, Chicago.
- Einstein, A., B. Podolski, and N. Rosen: 1935, 'Can Quantum Mechanical Description of Physical Reality Be Considered Complete?', *Physical Review* 47, 777-80. Reprinted in Wheeler and Zurek (1983).
- Fine, A.: 1981, 'Einstein's Critique of Quantum Theory: The Roots and Significance of EPR', in Barker and Shugart (1981), pp. 147-58.
- Fine, A. and P. Machamer (eds.): 1986, *PSA 1986*, Philosophy of Science Association, East Lansing, Mich.
- Gasiorowicz, S.: 1974, *Quantum Physics*, Wiley and Sons, New York.
- Glymour, C.: 1980, *Theory and Evidence*, Princeton University Press, Princeton.
- Glymour, C.: 1984, 'Explanation and Realism', in Leplin (1984), pp. 173-92.
- Harman, G.: 1965, 'The Inference to the Best Explanation', *Philosophical Review* 74, 88-95.
- Harman, G.: 1968, 'Knowledge, Inference, and Explanation', *American Philosophical Quarterly* 5, 164-73.
- Humphreys, P.: 1989a, *The Chances of Explanation*, Princeton University Press, Princeton.
- Humphreys, P.: 1989b, 'Scientific Explanation: The Causes. Some of the Causes, and Nothing But the Causes', in Kitcher and Salmon (1989), pp. 283-306.
- Kitcher, P.: 1989, 'Explanatory Unification and the Causal Structure of the World', in Kitcher and Salmon (1989), pp. 410-506.
- Kitcher, P. and W. C. Salmon (eds.): 1989, *Scientific Explanation*, University of Minnesota Press, Minneapolis.

- Kochen, S. and E. P. Specker: 1967, 'The Problem of Hidden Variables in Quantum Mechanics', *Journal of Mathematics and Mechanics* 17, 59-87.
- Leplin, J. (ed.): 1984, *Scientific Realism*, University of California Press, Berkeley and Los Angeles.
- Locke, J.: 1975, *An Essay Concerning Human Understanding*, (first published in 1689) edited by P. H. Nidditch, Oxford University Press, Oxford.
- Mayo, D.: 1986, 'Cartwright, Causality, and Coincidence' in Fine and Machamer (1986).
- Perrin, J.: 1923, *Atoms*, translation of *Les Atomes* (1912), translated by D. L. Hammick, Van Nostrand, New York.
- Railton, P.: 1989, 'Explanation and Metaphysical Controversy', in Kitcher and Salmon (1989), pp. 220-252.
- Salmon, W.: 1984, *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton.
- Salmon, W.: 1989, 'Four Decades of Scientific Explanation', in Kitcher and Salmon (1989), pp. 3-219.
- Schilpp, P. A., (ed.): 1949, *Albert Einstein: Philosopher-Scientist*, The Library of Living Philosophers, Evanston IL.
- Van Fraassen, B. C.: 1980, *The Scientific Image*, Oxford University Press, Oxford.
- Van Fraassen, B. C.: 1984, 'Belief and the Will', *Journal of Philosophy* 81, 235-56.
- Van Fraassen, B. C.: 1985, 'Empiricism in the Philosophy of Science', in Churchland and Hooker (1985), pp. 245-308.
- Van Fraassen, B. C.: 1989, *Laws and Symmetry*, Oxford University Press, Oxford.
- Wheeler, J. A. and W. H. Zurek (eds.): 1983, *Quantum Theory and Measurement*, Princeton University Press, Princeton.

Manuscript submitted September 19, 1990

Dept. of Philosophy
1001 Cathedral of Learning
University of Pittsburgh
Pittsburgh, Pennsylvania 15260
USA

PROCESS CAUSALITY AND ASYMMETRY*

ABSTRACT. Process theories of causality seek to explicate causality as a property of individual causal processes. This paper examines the capacity of such theories to account for the asymmetry of causation. Three types of theories of asymmetry are discussed; the subjective, the temporal, and the physical, the third of these being the preferred approach. Asymmetric features of the world, namely the entropic and Kaon arrows, are considered as possible sources of causal asymmetry and a physical theory of asymmetry is subsequently developed with special reference to the questions of objectivity and backwards causation.

It is commonly held that the relation between cause and effect is asymmetric. That is to say, it does not follow from 'x caused y' that 'y caused x', for there is a fundamental distinction between 'cause' and 'effect': causation has directionality. This question of directionality or asymmetry has long been an important ingredient in philosophical discussions of causality. We should say, at the least, that an adequate theory of causality should provide an explanation or account of causal asymmetry. This paper looks at the capacity of 'process' theories of causality to provide such an account.

Ducasse (1975), Russell (1948), Aronson (1971), Fair (1979), Salmon (1984) and Dowe (1992) have all contributed to the programme of understanding causality as a property of individual processes. Of these only the latter three theories will be discussed in detail. Broadly speaking, these theories take (continuous) processes as more fundamental than events, and seek to identify what properties a process might have which would qualify it to be a causal process. For Fair it is energy; for Salmon, marks; and for Dowe it is conserved quantities. Now energy transference, marks transmission, and the ascription of conserved quantities are all reversible phenomena, governed by the time symmetric laws of nature. We will therefore need to consider how these theories account for causal asymmetry. In doing so there are three general criteria or constraints that I would like to obtain in an account of asymmetry:

Erkenntnis 37: 179-196, 1992.

© 1992 Kluwer Academic Publishers. Printed in the Netherlands.