SOME REMARKS ON THE NOTION OF CAUSALITY

J. B. CRANSTON

Aristotle tells us that to do science is to discover what a thing really is, to learn the thing's reason for being. This requires that the essential cause or causes of the thing be undiscovered. It requires telling what the thing is, that it is, why it is and if it is. In other words, to do science is to explain events and objects in terms of their causes.[1]

Since Galileo's time, science is thought to have become quite a different sort of enterprise from that discussed by Aristotle. It has become empirical. It has been mathematized. Above all, it has been mechanized. Post-aristotelean science deals with inanimate objects whose behavior can be described by mathematical formulas. Gone is the animistic world of aristotelean science with its notions of natural place and self-moved bodies with souls, its universe of concentric spheres and its separate laws for celestial and terrestial realms. Final cause has been downgraded from the ultimate reason for a thing's being to a mere anthropomorphism. Henceforth, physics would deal with dead matter and formal mathematical laws.[2]

Nevertheless, post-scientific revolution science was still deeply entangled in the notion of causality. Aristotle's definition still applied. Science still involved, in an essential way, the attempt to explain phenomena in terms of causes. Only now, the four causes of Aristotle had been reduced to one: efficient cause. But the new mathematical laws of nature were still causal laws. They gave the reason for a thing's being, or at least, for its behavior. Galileo had temporarily set aside the search for an explanation of the ultimate nature of things, in the belief that it is first necessary to describe and understand their physical behavior. Science has yet to return to the former problem. Thus, a more limited causality was applied in classical (Newtonian) physics, a causality that could be expressed in precise mathematical equations and which dealt only with the physical states (motion and position) of inanimate bodies.[3]

The first thirty years of the twentieth century brought a new revolution in physics, a revolution which has had grave consequences for the concept of causality. Quantum theory has introduced a fundamental uncertainty with respect to some of our underlying physical concepts, including, especially, causality. And this uncertainty does not emanate from purely philosophical considerations, but is based in scientific theory and supported by empirical observation. The mathematical equations of quantum theory do not permit one to describe the behavior of micro-particles in ordinary, familiar terms of continuous paths between well-defined points

in space-time. The quantum of action injects an atomic or discontinuous aspect into the concepts of momentum and energy—at least at the sub-atomic level. Atomism is no longer restricted to the realm of matter but has been extended and generalized to include the dynamic laws themselves. Thus, the laws which describe and explain the behavior of sub-atomic nature appear to many to be non-causal. At the most fundamental level, causality seems to break down. And the causality which seems so apparent in the every day world of directly sensible objects may turn out to be no more than an appearance, an illusion resulting from the gross level at which sense experience occurs, a measure of our inability to distinguish with our senses the finer structure of nature.

For a number of reasons which it would be inappropriate to discuss here, it seems to me that the notion of causality is an important, if not essential, part of the scientist's metaphysical bag (as well as of his epistemological bag). I'm not sure that he could get along very well without assuming some sort of causal connection between the various phenomena he describes. The connection between himself and what he is observing is particularly crucial. For this reason, I shall attempt here to salvage causality.

The first thing to do, I suppose, is to determine what is meant by 'cause' or 'causality' as the concept is applied to the physical world. Classical mechanics is often held up as the paradigmatic case of a causal system. So, perhaps a look at it will be enlightening.

One of the crucial characteristics of classical mechanics, and of the mechanistic paradigm of which it forms the heart, is its predictive power. Given certain information about the present state of affairs in a physical system, future states of affairs may be predicted with a high degree of accuracy and certainty. In principle, absolute accuracy and certainty are possible. More precisely, given the physical state of a system (e.g., the position and momentum of a body) at time t_0, and given the absence of interference from outside the system, the laws of physics will yield the physical state for any future time t_1—or for any past time, for that matter. It is assumed, of course, that the laws of physics are invariant with respect to time and space. This forms the basis of Laplace's famous intelligent being argument in which he extends these ideas to a system which takes in the whole of the universe.

Systems of this type are called deterministic, and classical physics, to the extent that it actually conforms to this model, is deterministic. Such systems are also referred to as causal systems. Thus, in classical physics there is a strong tendency to equate causality with determinism. But, is this justified?

Another way to grasp what is meant by 'cause' is to look at how the concept is used in a somewhat broader context. It turns out that it is possible to identify two basic senses of cause in ordinary language. One I will call the agent-action sense and the other the explanatory or descriptive sense. Suppose I am asked, "What caused the baby to fall?" I answer, "His brother pushed him." That is, I cite an agent-action as the cause. Countless similar examples can be given, involving either living (intentional) or inanimate agents. "What caused his eye to be blackened?" "The ball hit him in the eye." The ball is the agent here. We conclude, then, that sometimes to ask for the cause of something is to ask that an agent-action be named.

On the other hand, suppose I ask, "What caused the water in the pot to boil?" I might accept an agent-action type of answer: "Mary turned on the burner." But, then again, this might not be the kind of answer I am looking for. I might want to know the *cause* of the water boiling, i.e., a more general kind of answer, an explanation of why (what causes) water to boil when heat is applied. I am still looking for a cause, but now a simple agent-action answer won't do. An answer that would satisfy me would have to do with the transfer of energy, the laws of thermodynamics, etc. The question asks that the phenomenon of water boiling be accounted for, the phenomenon in general. It asks for an *explanation*, or perhaps a description, of what happens in such cases. This is the other sense of cause: causal explanation.

As is evident in the case of the boiling pot, the agent-action sense and the explanatory sense often overlap. Basically to ask for the cause of something is to ask that it be *accounted for*. "Accounting for" is common to both senses. But the explanatory sense is the more general, the abstract and scientifically relevant, sense. Simply naming an agent-action is often inappropriate or does not adequately account for the event. Unlike Aristotle, modern scientists don't talk much about causality. (They simply assume it.) But science is still involved in trying to *account for* our experience of the physical world. It tries to give the reasons for physical events and phenomena, to explain them. In science, when one asks that a phenomena be accounted for, one is not asking that some causal agent be identified but that the phenomena be related to, or interpreted in terms of, some system of descriptive laws and theories. One is asking for a causal explanation. How did science account for the spectrum of the hydrogen atom? Early quantum theory was created, in part, as an explanation of this phenomena—to give its cause. The theory both discovers the cause and, in an epistemological sense, is the cause. David Bohm tells us in his *Causality and Chance in Modern Physics*[4] that to ask for the cause of a phenomenon is to ask for an explanation of the facts. To give a causal explanation is to "fit" the phenomenon into a body of theory and law, a process which may involve some adjustment to the body of theory itself.

If causality is an accounting for, i.e., a causal explanation, then what is its relationship to determinism? And what is the impact of this point of view on the problem presented by quantum theory?

The classical mechanistic system discussed earlier is clearly a causal system. It provides causal explanations for certain classes of physical phenomena. As we have seen, it is also a deterministic system. Thus, at least some causal explanations are deterministic explantions. Now, deterministic explanations are so constructed that each event uniquely determines another event and so that each event is itself in like manner uniquely determined. Nothing could have happened otherwise. Furthermore, deterministic systems are reversible. There is no ambiguity with respect to the past just as there is none with respect to the future. The question arises, might a physical explanation which lacked one or both these features, i.e., uniqueness of determination and reversibility, still count as a causal explanation? It seems to me that the answer is yes.

A causal explanation is an explanation which *accounts for* the phenomena. Clearly, there is no intrinsic requirement that the explanation be reversible, that it be simultaneously predictive and postdictive. The laws of classical mechanics possess this characteristic, but those of thermodynamics do not. Does this mean that thermodynamics fails as a causal explanation? On the primitive, intuitive level causality is undirectional, just as is heat theory. If we accept thermodynamics as a causal explanation, then we have already made a distinction between causality and determinism.

However, the distinction may be pushed further. A causal explanation need not *uniquely determine* at all. That is, the relationship between physical events within a causal system need not be a one-to-one relationship. Within the context of a given set of circumstances, a particular physical event may produce any one of several future events. A single cause may determine a set or a range of possible effects rather than a single, unique effect. And this is in accord with everday experience. Fix a rifle in a vice in such a manner that it is correctly aimed at the bull's-eye of a distant target. Fire the rifle repeatedly. What will be the outcome? A pattern of hits clustered about the bull's-eye. The best that the appropriate physical theory, i.e., classical mechanics plus error theory, can predict for a single hit is that it will fall within a certain area. (Probabilities for various portions of the area may be established also.)[5] One causal event, numerous possible effects. In a similar manner, a single event may be determined by a range of causes.

It might be objected that this line of reasoning is faulty in that it overlooks an essential aspect of the problem. The non-deterministic causal relations suggested above are actually apparent rather than real, the result of ignorance with respect to certain causal factors relevant to the situation. If one had available sufficient information, so the argument goes, then one would see that the outcome is actually totally determined and that only one result was ever really possible. The correct causal explanation is the one which takes into account *all* relevant circumstances. In the case of the rifle this would include temperature, air currents, chemical composition of the powder, weight and exact shape of the projectile, etc., etc.

In response to this, I would simply point out that *any* possible explanation must be finite with respect to the circumstances it takes into account.[6] This is the nature of a physical explanation. It is necessarily finite, both quantitatively and qualitatively. To meet the requirements of the determinist would ultimately require consideration of the totality of the universe, an impossible and meaningless project. On the more mundane level, would the determinist extend the relevant circumstances in the rifle example to include the molecular state of the matter making up the projectile? If this factor is excluded, the prediction will be less accurate. That is, the prediction will not be absolute, it won't be deterministic. Both in practice and in principle, a relevancy line must be drawn somewhere, and wherever it is drawn it will exclude some factors, possibly infinite in number, which may affect the outcome. Relative to the factors ignored or excluded, there will be an element of indeterminism or chance, if you will, in our causality. Determinism works as an explanation in many realms simply because these are realms in which those factors which have been excluded can be safely ignored for purposes of the problem at hand. Deeper analysis will always raise an element of indeterminism with respect to broader contexts. Determinism, it appears, is but an abstraction from the more general concept of causality.

It would seem, then, that causality is not to be equated with determinism. As a matter of fact, causality appears to be compatible with least some form of indeterminism. However, there are two aspects of causality which tend to moderate its association with indeterminism. These are the notion of continuity and the assumption that every physical event has a cause. Both appear to be vestiges of a primitive, perceptual notion of cause and effect which have been carried over into the abstract concept of causality. Nevertheless, causal explanation must take them into account. That is, a causal explanation must be able to provide an account of events which shows each event as a consequence of other events and which presents no event as unexplainable, i.e., as uncaused. These requirements are two sides of the same coin. To show causal continuity and to show that no event is uncaused amount to the same thing. It is to show that the physical explanation is adequate: that it is internally coherent, that it is

compatible with the existing theory of which it forms a part (or that existing theory must be altered to accommodate it), and that it accounts for the facts (i.e., the phenomena). However, it is not necessary to show that the explanation is deterministic, that events are uniquely related in a one-to-one manner.

A deterministic explanation, of course, meets these requirements. But so do certain types of non-deterministic explanations. If an explanation is such that, given one event, the range of possible consequent events is specified, then the explanation is causal. Probabilistic or statistical explanations are of this type.[7] They describe the behavior of physical entities in terms of the statistical or average behavior of the group. This can be viewed as a case in which one event may produce any one of a number of subsequent events. These are explanations which take into account the context of the explanation, which take into account that there are relevant factors which have not been, or can not be, fully considered. (Some statistical theories claim to be but practical tools for explaining the gross behavior of groups of physical entities, behavior which could in principle be explained deterministically in terms of the individual entities. The kinetic theory of gases is an example. But I suspect that deeper analysis of the problem may show it to be otherwise.) In any case, a probabilistic explanation admits an element of chance or indeterminism—but it does not admit lawlessness. Individual events are undetermined only within certain well-defined limits, while average behavior is well-defined, determined even.

The application of these ideas to quantum theory should be fairly obvious. Quantum theory accounts for only the average or statistical behavior of subatomic particles. The Schrödinger equation, for example, permits us to calculate the probability that a particle will be found at a certain location, should a measurement be made. It accounts for the event by describing a continuous evolution of probabilities. Alternatively, it can be understood as accounting for the average behavior of a swarm of particles. In either case, it can be viewed as constituting a non-deterministic causal explanation. Given a particular physical event at time t_0, it predicts the likelihood of various possible consequent events at time t_1. The fact the quantum theory is probabilistic is not incompatible with it being causal.

However, there still remains a difficulty. Quantum theory is not only probabilistic, it is *essentially* probabilistic. That is, the statistical behavior of elementary particles *in principle* cannot be reduced to a deterministic explanation. Quantum physical states can change only by some multiple of the quantum of action. If the values of states A and B represent consecutive multiples of the quantum of action, then a particle is permitted to be in state A or in state B but it cannot have a state value between these of A and B. The particle can, under proper conditions, make a transition from A to B, but it must do so without ever being "between" the two. The theory can't account for the "path" between states. In fact the theory seems to deny the existence of such "paths." Thus, as a causal explanation, even as a non-deterministic causal explanation, quantum theory seems to have serious problems.

On the other hand, quantum theory does *account for* quantum events, even though there seem to be "holes" in the account. And perhaps these "holes" are not as serious as they at first appear. The concept of non-deterministic causal explanation is founded on the idea that any causal explanation is context limited and that factors not within the context, if they turn out to be relevant, will appear as chance factors relative to the explanation. The fact that one and only one of several possible outcomes actually occurs in a given test raises a problem for probability theory itself, and for non-deterministic causality in general, which is similar to the quantum "path" problem. The theory cannot account for the "choice" itself, i.e., for how the actual selection of the outcome is made. This question involves factors external to the context of probability theory. Yet, the theory functions as an explanation. Perhaps the situation is analogous for quantum theory and the problem of "path."

On the basis of this kind of reasoning, I conclude that quantum theory should count as an attempt at constructing a causal explanation or, perhaps, as an approximation to such an explanation. Despite Bohr and the Copenhagen interpretation, it seems to me that the theory, as it has been developed thus far, is basically causal in nature. (This is not to deny the radical nature of the transition from classical to contemporary physics.) At the very least, it has not been demonstrated that micro-nature is fundamentally acausal. What has been demonstrated is only that deterministic explanations can not account for micro-phenomena. It should not surprise us that some aspects of nature do not follow the deterministic model. But, it would be truly astounding if nature turned out to be acausal at its most fundamental levels.

NOTES

[1] *Posterior Analytics*, II. 1-2.

[2] One has only to look at quantum physics to realize that this view is no longer adequate. Actually, neither final causality nor animism ever completely died. They have coexisted along with mechanism ever since the scientific revolution, e.g., in the important concept of least action.

[3] This is not strictly correct. In the wave theory of light, a different notion of "physical state" and, I would argue, a different idea of causality were involved. However, the model was nevertheless mechanical and mathematical.

[4] David Bohm, *Causality and Chance in Modern Physics* (Philadelphia: University of Pennsylvania Press, 1971), p. 12. This thesis underlies this book.

[5] I am indebted to David Bohm for this example. *Causality and Chance in Modern Physics*, p. 16.

[6] This thesis, which is fundamental to my argument, is based originally on Bohm's notion of the qualitative infinity of nature. Its application and development here is, of course, mine, not Bohm's.

[7] Paul Bernays discusses this kind of causality in his brief but interesting article, "Causality, Determinism and Probability," *Perspectives in Quantum Theory*, eds. Wolfgang Yourgrau and Alwyn van der Merwe (Cambridge: MIT Press, 1971), p. 260.

University of New Mexico