Philosophy of Science // Fall 2016

Handout 12

Mechanistic explanation: Wright and Bechtel, Craver

The possibility of ontic explanation. In the D-N account of explanation the explanatory job was done by a logical inference from laws and initial conditions to the explanandum. In this sense the D-N account is *epistemic*: what explains is a certain inference we have formulated. That is, in explaining a phenomenon we come up with a logical inference. What is more, the laws themselves featuring among the premisses of that inference are regularities identified by *us*. Thus explanation becomes entirely a product of our own activities.

The objection here is that an adequate explanation must have a root in reality. There can be all sorts of 'convenient' explanations giving us a sense of mental satisfaction, a sense of understanding, and a D-N model can give us that. Yet an adequate, 'true' explanation should in addition track a feature of reality. An adequate explanation must by necessity be ontic: it must exhibit a feature of reality among its explanans.

What could that feature be? One proposal is, causes. An adequate explanation must exhibit a causal relation leading up to the occurrence of the explanandum. It must, that is, give us a causal mechanism behind the phenomenon (see the telling quotations in 47–48 of Wright and Bechtel).

This contrast is problematic for a number of reasons. In the first place, why are laws not features of reality? Laws cannot be contained in space and time like chairs and tables, but they can still objectively characterize reality. We could then say that we discover laws, not invent them.

This rejoinder is not altogether successful. For one, it presupposes a nomological necessity. We say that things happen in the world in virtue of its laws. This move may be promising, but not for the original proponents of the D-N view. According to them, laws were mere regularities, and necessity was banished. The whole motivation behind the deductivist approach was to dispense with necessity in the first place. Secondly, we have lost the appeal to logical inference along the way. We now say that, for example, the Moon orbits the Earth *because* of the law of gravitation, not merely because there is a logical inference from the law of gravitation to its motion. Such inference can help us illustrate the regularities of the planetary motion. But it does not itself perform the explanation.

But the contrast still appears poorly drawn. For consider: items in the world, such as mechanisms, do not themselves explain. They are simply not the kind of things that explain, because in explaining we have to arrange them in a certain way, describe them—in short, represent them. Suppose we work with a causal mechanism, such as a lighter. The operation of the lighter explains fire. But even in order to attribute causality to the lighter, to identify it there, we must have a theory about how lighters operate. Our explanation will then involve a description of a mechanism, not the mechanism itself. Every explanation appears to involve an epistemic element, a fragment of human activity.

We may still be able to contrast mechanistic explanation with a nomological one if we insist on the representational power of mechanisms. What explains, according to this thought, is our description of a mechanism, rather than our representation of laws and the associated D-N inferences. Laws can play an auxiliary role in our descriptions of mechanisms. However, the ultimate explanans is the mechanism—or rather, a representation thereof.

Question 1. In what sense does a D-N explanation count as an epistemic one?

What is a mechanism? At this point we have to ask what mechanism actually is. The general idea is that it is a complex system consisting of parts. But what makes them parts *of* the mechanism? Here is one proposal:

> A machine (=mechanism) is a composite of interrelated parts, each performing its own functions, that are combined in such a way that each contributes to producing a behavior of the system (Wright and Bechtel, 44).

This is hardly satisfactory. Is, on this account, the planet Earth a mechanism? We do not get a good answer unless we say what relates those parts and how they contribute to the system. Everything is related to everything else, and everything contributes to the behaviour of everything else—unless we are more specific about the nature of those relations and contributions.

A more refined suggestion is that these parts should be assembled in a hierarchy. But a hierarchy of what? To rule out absurd alternatives, we now say that this hierarchy simply reflects causal relations. But then the fancy talk of mechanisms has no theoretical role to play. It becomes merely a fragment of a causal network.

The fact that mechanisms consist of parts allows us to say that mechanistic explanations are constitutive. They explain by reference to the mechanistic parts, the components of a given mechanism. But if these parts are related by causation alone, then the constitutive explanation becomes a species of causal explanation.

Are parts real, or are they just heuristic devices allowing us to talk intelligibly about the mechanism? Craver gives us five criteria for their reality: stable properties, robustness, possibility of intervention, theoretical plausibility, and relevance. Of these, robustness and intervention are most significant.

Laws reconsidered. Craver lists five desiderata of mechanistic explanation. It must list the conditions under which the phenomenon (the output of the mechanism, so to speak) actually occurs, under which it fails to occur, it must list the influences that can change the character of the phenomenon, it must describe the behaviour of the mechanism under non-standard conditions, and it must list by-products of the mechanism's performance. All of these are plausible demands, though it is not clear why it should be those demands exactly. Still, these demands can only be met if we are in possession of laws governing the behaviour of the mechanism's parts. Until and unless we have those laws, all that would simply be *another description*.

Models. This last point is reinforced by the discussion of models. Briefly, the idea is that models should describe counterfactual behaviour of mechanisms. Mere generalisations available in 'phenomenal' model are not explanatory, even though they can useful in prediction. But in order to create such models, you have to have an idea of laws. What remains to be a problem is the distinction 'real' laws and inductive generalisations, familiar to us from Goodman.

YSB