14.D: the ***Philosophyldeas*** overview of

Explanation

The truth is no use if it is too complex for us, and prediction is frustrating if we don't know its cause. What we really want is **understanding**, and explanation seems to be our means of achieving it. Hence explanation is central to science and philosophy, and a concept which needs close study. Explanation is a particular challenge to empiricists, because it offers understanding, but often by speculatively reaching beyond the patterns of ordinary experience. Instrumentalists, who just confine knowledge to available data, have little faith in so-called 'explanations'.

The initial questions about explanation are 'where does it begin?' and 'where does it end?'. Explanation is a human activity, and **begins** for most people when they meet surprises and irregularities, but curious people (philosophers!) also want to explain what is normal and taken for granted. The best word to cover all of this is '**puzzles**'. However, this makes explanation relative to puzzled persons, and degrees of puzzlement. What is obvious to one person is a puzzle to another. If explanation is just what removes a puzzle, a good explanation might just 'fob you off' instead of being illuminating. We need to talk of 'genuine' puzzles, and 'genuine' explanations, which sounds a bit vague.

It is a notorious difficulty when we ask where explanation **ends**, that even the simplest explanation (if it is to be complete) goes right back to the Big Bang. We could respond by saying it ends when a puzzle is dissolved, but this needs to be a 'genuine' solution, and not one that ends when the puzzled person gets bored. We don't actually expect explanations to return to the Big Bang, so puzzles must be **localised** in some way. Maybe the puzzled person decides the locality of their puzzle, but there may also be natural localities for explanations. We don't expect normal explanations about chemistry, for example, to wander off into physics or biology.

If most explanations start with genuine puzzles, and end when they have offered a story confined to the right locality, then we must ask *how* we should go about giving good explanations. This invites a prior question, which asks what features of reality can be drawn on with which to construct explanations. If reality is just a 'flat' array of experienced fragments, then puzzles will be patterns in the fragments, and explanations will be further patterns connected to them. A more fertile picture for lovers of explanation is reality having a '**structure**', so that the hidden underpinnings of a puzzle can be revealed and mapped. Thus the theory of explanation is usually given in terms of grounding, dependency, causation and determinations. Tracking this structure may lead down to 'ultimate' explanations, but shallow explanations are useful, and explaining a puzzle by a further puzzle can still be revealing.

An interesting phenomenon here is the '**direction**' of explanation. If a flagpole has a shadow, then it seems that the flagpole explains the shadow, but the shadow cannot explain the flagpole; a coming storm explains a barometer reading, but barometers don't explain storms. Presumably this tracks the direction of causation, which invites the idea that all explanation is merely the tracking of causes. However, the world contains a lot of causes, but we are selective about which causes make good explanations. If we want to confine an explanation to a locality (rather than tracking causes endlessly backwards), we can see how causes bunch together, and converge at certain points. Explanations that identify this phenomenon will be 'fruitful', which is a virtue of powerful explanations. The theory of evolution is famous because it offers to explain so much by identifying one specific aspect of causation in the processes of life.

Explanation that relies on focal points in the network of causation will tend to emphasise '**mechanisms**', and many famous explanations, such as the plate tectonic account of earthquakes, do just that. But the most famous explanations in the fundamental sciences, such as those of relativity and quantum mechanics, are usually presented to us as mathematical **equations**. Since equations and graphs just map correlations, we may wonder how far they qualify as 'explanations', rather than mere devices for describing and predicting. They do, however, reveal unexpected links (such as the gravitational link between apples and planets), which certainly enhances understanding.

Ancient thought identified four or five '**modes**' of explanation. These modes of explanation may complement one another, without being directly connected. A complete explanation would need all of the modes. The idea was to give the 'material' aspect (what constitutes the thing), the 'efficient' aspect (that triggers it), the 'formal' aspect (its concepts and principles), and the 'teleological' aspect (its purpose). A fifth mode is the revelation of the **necessity** of the thing. Thus if you show why a road accident *had* to occur, your explanation is fully successful.

If an explanation reveals the necessity of something, that invites the idea that logical deduction is involved, and one suggestion is that the explanation of an event is a background situation plus a '**covering law**', from which the event can be logically inferred. This is the cautious empiricist account of explanation, which fits an event into our pattern of experiences. If we understand the laws of nature, then showing which law is operating in the event seems a good route to understanding. However, this logical approach fails to capture the direction we find in explanations, and laws are idealisations, whereas real events are a tangle of intersecting laws. If laws are regularities in experience then a lawlike explanation just says an event fits a regularity, and that doesn't seem to explain anything.

An important strategy proposed for explanation is to introduce **contrasts**. If we just say 'explain this event', with no mention of what aspect concerns us, it is hard to know where to start. If we say 'explain why this event rather than that one', an explanation can be much more focused, while remaining more objective than a response to a personal puzzle about the event.

The biggest sceptical doubt about explanations is that people love explanations too much, and will grab at an explanation even when remaining in doubt is the better option. Inference to the **Best Explanation** is an approach to reasoning about the world that has many supporters, but the obvious problem is that the best explanation may still be a wrong explanation. If we can identify which explanation is the 'best', we still need further criteria to decide whether to accept this best candidate. The explanatory virtues will include simplicity, generality, convergence on a source, coherence with other explanations, and fruitfulness in extending understanding. Optimists hope this will lead to 'real' explanations, but a more cautious view is that good explanations are just useful.