

Chapter 8

Explanations

'After the professor's explanation we were just as confused as before, only now on a higher level.'

8.1 Explanation and Prediction

The goal of science is a much-debated issue within philosophy of science. Is the goal of science to produce scientific explanations, or is it to make predictions? Some people attribute an essentially psychological meaning to the word 'explanation'; and hence infer that explanations are not important for the sciences. Instead, it is claimed that what is essential to science is accurate prediction. On the other hand, some say that the central aim of science is to satisfy our curiosity in regards to how the world works; to understand physical processes, social processes, man's history and people's reasons for their actions, which is acquired through explanation. Thus explanation is essential to science, whereas predictions only have to do with the application of science in technology. Technology is important, of course, but it is by-product of our understanding of how the world works.

A different view of this matter, held by Hempel, among others, is that the only difference between explanation and prediction is the temporal order of the act of explanation or prediction and the relevant event itself. Thus explanations regard past events, whereas predictions are about future events, but the relation between the *explanans* (that which explains/predicts) and the *explanandum* (that which is to be explained/predicted) is the same. Therefore, according to this view there is no interesting difference between explanation and prediction. It follows that if one can predict future events, then one can thereby also give a good explanation of those events once they have occurred.

I don't think Hempel's view holds water. There are many examples where one can make good predictions, but those predictions are not considered good explanations of a certain phenomenon. One can also give examples of the reverse: situations where good explanations can be given, and yet no relevant predictions can be made on the basis of those explanations. As an example of the first case, consider the correlation between the motion of the moon and ocean tides. Many people have observed this correlation over millennia; particularly in England,

where tidal ranges are large and important for navigation at sea ports. High tide follows only a few hours after the moon travels past the meridian. Given that one has been able to rather accurately calculate the moon's motion ever since the time of the Babylonians, it has long been possible to predict the tides with reasonable accuracy. However, there was no explanation for this correlation until Newton published his theory of gravity. The fact is that Newton's ability to explain the change in tides was considered a monumental step forward for science, even though one had previously been able to accurately predict high and low tides.

An example of the second case, that one can give a good scientific explanation of some phenomenon without the possibility of making a prediction, is found within biology. Many biological phenomena can be explained with the help of the two basic principles of evolution,¹ but these principles cannot be used to predict anything. For example, consider the long neck of the giraffe. It is not difficult to explain why giraffes have such long necks given that (i) they live in the savannah, and (ii) gene variation occurs randomly. The giraffes that, due to random genetic variation, happened to have longer necks than other giraffes could eat leaves higher up in the trees and hence were better fit to reproduce. After a number of generations, the entire population came to have long necks. However, from this mechanism one cannot derive any prediction of which traits giraffes will acquire in the future, since the mutation of giraffe genes occurs at random and the disadvantages of a very long neck may come to outweigh the advantages at some point in the future via habitat change or some other mechanism.

In what follows, I will take for granted that explanations are different from predictions, even though predictions are sometimes facilitated by scientific explanation.

8.2 What Is Explained?

In everyday situations the word 'explanation' is used in many different contexts. In university education, as in other formal educations, new concepts are introduced and explained. This can be done in different ways, such as by giving explicit definitions, or by giving examples that show how a certain concept is used. However, this is not the meaning of explanation as it is discussed in the philosophy of science. Rather, what we are interested in is scientific explanation.

A scientific explanation can be roughly characterized as a line of reasoning that satisfies our curiosity about a phenomenon. In fact, much research is driven by the researcher's desire to understand mysterious phenomena. According to this view, which I share, explanation is the goal of all scientific disciplines. The question that

¹ These principles are the following: (i) individuals of the same species struggle for survival and the chance to reproduce, and those best adapted to their environment produce more offspring, and (ii) genetic variation in a species occurs randomly.

then arises is whether explanations have the same form in every scientific discipline.

According to a well-known position formulated by Dilthey and Weber, among others, there is an essential difference between explanation in the natural sciences and explanation in the humanities/social sciences (German *Kulturwissenschaften*). According to Dilthey and Weber, explanations in the natural sciences aim at producing descriptions of causal mechanisms, whereas in the cultural sciences, explanations aim to give *understanding*. The concept of understanding, as it is used here, contains an element of empathy (German *Einfühlung*). The objects of such explanation are actions and the artefacts and institutions those actions produce; these actions and artefacts have an intentional component that require interpretation, and are therefore not suitable for causal explanation.

Contrary to Dilthey's and Weber's views, Carl Hempel² claimed that all explanation, regardless of domain, has a common structure. This thesis has been much debated.

Another overarching issue regarding explanation is whether it is possible to determine the correctness or force of an explanation independent of the context in which it is provided. Many philosophers take it for granted, almost unconsciously, that the explanatory force of a scientific theory is a context-independent property that can be determined objectively once we find the correct general structure of explanation. In contrast to this view, a few philosophers, most notably Bas van Fraassen, have claimed that explanatory force is not something that a theory contains per se. According to this view, explanations are fundamentally dependent on their context; they are answers to questions people ask.

I will start by discussing the most well known model of explanation, the so-called Deductive-Nomological model (D-N model), developed by Hempel and Oppenheim, which claim universal applicability. Then I shall take up its statistical counterpart, the Inductive-Statistical model of explanation, which will be followed by a number of modifications of the D-N model. Following this discussion, I will present models for explanation that do not claim universal applicability, ending finally with a discussion of Bas van Fraassen's pragmatic model of explanation. This list does not exhaust all the various models of explanation. The debate surrounding this issue is quite lively, and as a result many ideas have been discussed. This shows that the concept of scientific explanation is multifaceted. The discussion will continue in the next chapter with various aspects of explanation that are especially relevant to the humanities and social sciences.

In the forthcoming discussion, two technical concepts will be important: 'that which explains', or *explanans* and 'that which is explained', or *explanandum*.

² Carl Hempel (1905–1998) was born in Germany and worked in USA.

8.3 The D-N Model

The modern philosophical debate about explanation began with Hempel and Oppenheim's seminal 1948 paper *Studies in the Logic of Explanation*, in which they present the Deductive-Nomological model of explanation. The central idea of this model is that an explanation is a *deductive inference* from one or more *scientific laws* (law = _{Greek} nomos) and *initial conditions*.

A classic example is Newton's explanation of why all bodies fall with the same acceleration, regardless of how heavy they are. According to an intuitive understanding of falling objects, it seems that the heavier an object is, the faster it falls, which was the general belief up until Galilei disproved it at the beginning of the seventeenth century.

Newton's explanation builds upon two fundamental laws; the law of gravity and Newton's second law, which states that a force (F) exerted on a body is equal to the body's mass (m) multiplied by its acceleration (a), i.e. $F = ma$. The law of gravity states that two bodies attract each other with a force proportional to the product of their masses and is inversely proportional to the square of the distance between them, as is given by the formula

$$F = G \frac{m_1 m_2}{r^2},$$

where G is the universal gravitational constant, which Newton calculated to be $6.67 \cdot 10^{-11} \text{ Nm}^2/\text{kg}^2$ by experimenting with pendulums. If we treat falling as the result of the gravitational force between the falling object and the earth, we can equate the force due to gravity and the force causing acceleration, which gives us

$$m_1 a = G \frac{m_1 m_2}{r^2},$$

where m_1 is the mass of the falling body, m_2 is the mass of Earth, and r is the distance between the centres of the two bodies. This equation can be reduced to

$$a = G \frac{m_2}{r^2}.$$

Thus the acceleration of a body is determined by the universal gravitational constant, the earth's mass, and the distance between the falling object and the earth. Newton showed that the distance between two bodies, in this kind of situation, is the distance between their centres, and thus that we can use the earth's radius and a measure of r . The earth's mass is $6.0 \cdot 10^{24} \text{ kg}$ and its radius is $6.4 \cdot 10^6 \text{ m}$. Plugging these values into the formula above gives us the acceleration, namely 9.8 m/s^2 . Thus we have a mathematical, i.e. logically valid, inference to the conclusion that free fall acceleration is constant and independent of the mass of the falling body.

Such an argument is, according to Hempel, a good scientific explanation of why free fall acceleration is constant for all bodies. The two essential elements of this explanation are (i) that the argument is a *logically valid inference* to the explanandum, and (ii) that the explanans contains at least one scientific law. Hempel presents these two properties as necessary and sufficient *formal* criteria for explanation. The only other requirement is that the propositions contained in a given explanation are true.

Schematically, we can organize the explanation of free fall acceleration in the following way:

Law 1:	$F = G \frac{m_1 m_2}{r^2}$
Law 2:	$F = ma$
Initial Conditions:	Earth's mass = $m_2 = 6.0 \cdot 10^{24}$ kg Radius of Earth = $r = 6.4 \cdot 10^6$ m
Conclusion:	$a = 9.8 \text{ m/s}^2$

Hempel argued that all explanations must fulfil these two formal criteria, and conversely that all arguments that fulfil these criteria, and whose premises are true, are explanations. Thus he claimed that all explanations, if they are truly explanations, have this structure, regardless of whether they pertain to physics, biology, history or any other science. Before we examine whether these criteria are actually valid in sciences like history, let us discuss whether this model is correct within its primary domain of application, the natural sciences.

It is easy to find lots of examples of arguments within the natural sciences that we immediately understand to be good explanations, and that have the aforementioned structure, but it is just as easy to find counterexamples.

8.3.1 Problems with the D-N Model

- (i) An explanation requires access to scientific laws. How have we come to know such laws? Well, quite often we have guessed or inferred them from the phenomena they are meant to explain. For example, one argument for believing the law of gravitation is the observed behaviour of bodies in free fall. But for an argument to count as an explanation, it is reasonable to expect acceptance of the explanans to be independent of the explanandum. If the explanans is to explain the explanandum, then our reasons for deeming the explanans to be true cannot involve the explanandum. Hempel's model thus requires that before the explanation is presented, we somehow have knowledge of the scientific laws involved and accept them as true. As we shall see in the next chapter, this is a big problem.
- (ii) Explanations are asymmetric even as regards the laws. This asymmetry is missing in the D-N model. Consider the following explanation:

Law: If stormy weather is coming, the barometer falls.³

Initial Condition: Stormy weather is coming.

Explanandum: The barometer falls.

This seems to be ok. If someone asks why the barometer fell, one would give an explanation very similar to the above example. The argument is intuitively acceptable and fulfils the two criteria for explanation, but compare it to the following:

Law: If the barometer falls, then stormy weather is coming.

Observation: The barometer fell.

Conclusion: Stormy weather is coming.

This example also fulfils the requirements of the D-N model (for argument's sake, let's assume that the law 'If the barometer falls, then stormy weather is coming' is at least approximately correct), but one cannot very well accept it as a correct scientific explanation for why stormy weather ensues. That the barometer falls does not explain why there is stormy weather; rather, the opposite is the case: the stormy weather explains why the barometer fell. There is asymmetry between explanans and explanandum; that is, if A is part of the explanation for B, then B cannot be part of the explanation for A.

If we generalize the above example, we see that the requirement of asymmetry entails that the laws used in D-N explanations must also be asymmetric, in the sense that the reasons for believing 'If A, then B' are not good reasons for believing 'If B, then A'. This in turn shows that a correlation cannot suffice as a premise/law in an explanation. For if there is a strict correlation between A and B, then we can express it in the canonical form 'A if and only if B'. If this is true (in the case of partial correlations we can replace/generalise the expression 'If A, then B' with 'The probability of B, given A' and vice versa), then we do not have asymmetry. This shows the correlations lack explanatory force.

Furthermore, not all sentences of the form 'if A, then B' for which there is empirical support are accepted as laws. What is more needed than empirical support? This question is discussed in Chap. 10.

(iii) The D-N model allows for the explanation of a single fact from a general statement. This is normally not considered a real explanation. Here is one example:

Law: All ravens are black.

Initial Condition: The observed object is a raven.

Conclusion: The observed object is black.

³ It is doubtful whether we really should call this a law, because (i) it is a qualitative statement, and (ii) there are exceptions. But Hempel and Oppenheim didn't mention any strictures of what to count as a law and their own example is similar in these respects. And the problem here under scrutiny doesn't depend on these aspects.

One can hardly accept this argument as an explanation for why the observed raven is black, since in such a case it is not the singular observation that requires explanation, but rather the fact that *all* ravens are black; we would normally not classify the observed object as a raven if it were not black. If someone were to observe a white raven, we would all want an explanation for why it is white and not black like all the rest, but in the case of observing another black raven, one would probably want to know why *all* ravens are black, and not just why the observed raven happens to be black.

This problem arises because of the assumption, held by proponents of the D-N model that it is primarily singular events we want explained. This is often not the case. We often require scientific explanations of repeatable phenomena, and not of unique events. (Of course, the opposite can be the case in regards to human action, but we shall discuss this later.) Thus the D-N model does not place sufficient requirements on what can be accepted as an explanation.

At least two ways of improving this model have been proposed. The first is to restrict the types of laws allowed in explanations to only *causal laws*. One would then be able to say that the D-N model is a model for causal explanations, which is what Hempel and Oppenheim probably thought, even though they made no distinction between causal laws and non-causal regularities. The other proposal is called ‘unification’. The idea is that a phenomenon can be explained if one can show that it is logically associated with other previously unrelated phenomena.

8.4 Causal Explanations

One way to improve the D-N model is to restrict the type of laws required for explanations to only causal laws. With this requirement, one could say that the barometer example does not refute the model, since ‘if the barometer falls, then stormy weather is coming’ is not a causal law, and thus the barometer example is not an explanation. On the other hand, it is reasonable to claim that stormy weather causes the barometer to fall, that is, if one includes low air pressure in the concept ‘stormy weather’.

We have already discussed the issue of how one distinguishes between causal laws and other laws in the previous chapter. We saw that all laws that relate variables to each other, whether they are deterministic or probabilistic, are correlations, and we have identified the conditions that must be satisfied for there being a causal connection between two variables. Alas, we are often unable to determine whether these requirements are fulfilled, but the point is that there is a distinction. Returning to the barometer example above, with the help of meteorological knowledge we can determine that the connection between stormy weather and a low barometer value is a correlation caused by a common factor; namely, low air pressure.

As we will see in Chap. 10, scientific laws have the logical form ‘For all x , if Ax , then Bx ’. Thus a precondition for a causal explanation is that ‘ Ax ’ describes a cause of the event Bx . (See the Appendix for a discussion of the logical form.)

In summary, one can say that a causal explanation is a logical inference from the cause to the effect based on a causal law.

8.5 Explanation as Unification

Another modification that has been proposed to improve the D-N model is *unification*. The central idea is that, though an explanandum is logically inferred from the explanans in agreement with the D-N model, the explanatory force of an argument does not primarily lie in this particular inference, but in the fact that many disparate phenomena can be similarly explained. This union is accomplished when the explanans not only explains a particular explanandum, but also others simultaneously. In other words, a proposed scientific law that only explains one type of empirical phenomena does not have much explanatory force. To illustrate this idea, once again consider Newton’s law of gravitation. If it could be used only to infer free fall acceleration, it would not be much of an explanation of such phenomena. For a ponderous person could ask, ‘what reasons do you have for believing in the law of gravitation?’ The answer would then be that the law of gravitation explains the constant acceleration during free fall, which makes the reasoning circular. However, we know that the law of gravitation can be used to explain a great many seemingly unrelated phenomena for example:

- Kepler’s three laws for planetary motion,
- The motion of pendulums,
- The tides,
- The orbits of comets,
- The earth’s flattening at the poles,
- The precession of the vernal equinox

This means that our reasons for believing the law of gravitation are not limited to the particular phenomenon we want to explain in the first place, but that there are other independent empirical observations that serve as evidence for the law of gravitation. It is this property of the law of gravitation that gives it its explanatory force; the law of gravitation unifies the description of a number of phenomena into a coherent theory. An added advantage of this model is that it apparently contains different degrees of explanatory force, because it seems natural to talk about degrees of unification.

This view of explanation avoids the three previously mentioned problems with the D-N model: (1) we get an explanation of how we come to believe the explanatory law, (2) *unification* is asymmetric and (3) trivial generalizations have no explanatory force. However, it is not easy to formulate this idea into a concrete model. In this author’s estimation the most natural model is that presented by

Michael Friedman.⁴ Unfortunately, this model has been shown to have serious problems. Another variation, put forth by Philip Kitcher,⁵ has attracted more interest. However, a deeper discussion of these theories' strengths and weaknesses falls outside of the scope of this book.

8.6 Statistical Explanations

Statistical explanations are explanations that contain probabilistic arguments. Hempel argued that there is a strong analogy between deductive-nomological and so-called *inductive-statistical* explanation. The difference is – according to Hempel – that instead of deterministic laws in the explanans, there are statistical laws. It follows that the explanandum cannot, in a strict sense, logically follow from the explanans. Rather, it follows with a certain probability.

Statistical laws have the form 'Event Y occurs with probability p , given that event X has occurred', which in mathematical notation becomes $P(Y|X) = p$. Thus the structure of explanation is schematically

$$\begin{array}{l} P(Y|X) = p \\ X \text{ has occurred} \\ \text{-----}p \\ Y \text{ occurs.} \end{array}$$

The statement that Y occurs is not a logical consequence of the premises, which is indicated by the dotted line.

The word 'event' should be understood in the mathematical-statistical sense: an object having a property is an event. According to Hempel, a statistical explanation is good if the probability p is close to 1. The model is called 'inductive' because we cannot logically deduce that Y will occur, even if X has occurred. But we do have statistical evidence to rely on, which can be formulated into statistical laws.

One central problem, which Hempel himself noticed and discussed, is that the probability of an event occurring seems to depend upon how we describe that event. An example is the following: Olga Svensson, 56, works as a public servant in Sweden. What is the probability that she will vote for the social democrats at the next election?

Given that a large number of middle aged women working in the public service sector in Sweden vote for social democrats, one might guess the probability to be rather high; say 60 %. If we now suppose that Olga is a doctor, and that it is widely known that doctors' political preference is more conservative, then this probability must be revised to a lower value. If we further suppose that Olga is a professor at a

⁴ Friedman, M. (1974). Explanation and scientific understanding. *Journal of Philosophy*, 71, 5–19.

⁵ Kitcher, P. (1981). Explanatory unification. *Philosophy of Science*, 48, 507–31.

medical institute, then one must further lower the probability that she will vote for social democrats, since this portion of doctors tend to be even more conservative. Finally, if we point out that she is a member of the social democratic party, then the probability changes radically in the opposite direction. So what is the actual probability that Olga votes for the social democrats?

This example illustrates the so-called *reference class problem*: the probability of some event depends on which reference class one thinks the event belongs to. The probability that Olga votes social democrat depends on how you describe her; that is, to which reference class she belongs. If we describe Olga as a member of the social democratic party, then we get a high probability, and therefore an explanation of why she voted social democrat. But if we describe her as a professor at a medical institute, we get a high probability that she will not vote social democrat, and thus an explanation of the complement event. This is not acceptable. One cannot reasonably say that one both has an explanation of some event and an explanation of the 'opposite' event.

Hempel proposed that one could solve this problem by requiring *maximal specificity* for the event to be explained. Maximal specificity is related to available knowledge. In the case above, the event X would be a complete description of all of Olga's *known* properties.

The problem with this solution is that the resulting inductive-statistical explanation is relative to what we know. Hence explanatory force cannot be an objective property of the theory. This conflicts with the fundamental assumption, which many philosophers – including Hempel – made early on, that scientific explanations must be objective and independent of the knowledge of observers. Given that explanatory force is not objective, Hempel was forced to give up the connection between explanatory force and truth.

It might be possible to avoid the relativism of explanation due to available knowledge by stating that maximal specificity includes descriptions of all properties that effect the probability of an event's occurrence, known or otherwise. Assume then that we have described a certain event completely; that is, all circumstances that influence the probability of an event's occurrence are taken into account. Is it not then likely that we will find that the probability of this event's occurrence is either zero or one, since we have all the relevant information necessary to make that determination? If so, then we no longer have a statistical explanation, but a deductive-nomological one. We thus may infer that there are no such things as objective inductive statistical explanations. Conversely, we should accept that where these explanations are employed, as they often are in the sciences, that the probabilities used in these explanations do not reflect objective, observer-independent facts.

The only way to sidestep this conclusion is to say that inductive-statistical explanations are only applicable to genuinely random events. A genuinely random event is such that even if we know everything there is to know about the conditions

for an event's occurrence, we still cannot predict what will happen with certainty. I'll return to this topic in a moment.

Another problem is that, in many cases, we accept a probabilistic argument as a good explanation for events that are believed to be not genuinely random even though the event in question's probability is quite low. The connection between lung cancer and smoking is good example. Suppose a person has contracted lung cancer and that we ask for an explanation. We then view the information that the cancer victim smoked more than one pack of cigarettes per day for over 30 years as a good explanation. But, in fact, the probability of getting lung cancer if you smoke for over 30 years is only about 6 %. Still, we say that smoking is a good explanation for lung cancer. Therefore, Hempel's requirement that the probability should be high cannot be generally reasonable. Rather, it seems that the *increase* in probability when given new information is the essential explanatory step. The probability that a person will get lung cancer is very, very low, but if we add the information that a person has smoked for over 30 years, the probability increases many times over from this low baseline. It is this relative increase that we see as explanatory. This view suggests that what we take to be a good explanation depends upon what we know before the event, as well as what we find out from the explanation. Furthermore, this example and the entire discussion about inductive-statistical explanations also suggests that explanations have a contextual component that cannot be neglected. I will return to this point in Sect. 8.8.

I mentioned earlier that one could avoid the reference class problem in cases where the explanandum is a genuinely random event.⁶ It is easy to see that this is the case. Suppose we describe an event in a certain way and give the probability of its occurrence. Suppose further that all relevant circumstances are contained in this description. This means that the addition of any other information will not change the probability. Thus the probability of the event's occurrence is independent of any further specifications, and a complete description of the event, if it were possible to give one, would also not change the probability. But then this probability must be the result of a genuinely random event, and is not due to incomplete knowledge of relevant circumstances.

Of course, one can never be certain that a probability is the result of genuine randomness. However, according to our best present theories about nature, there are some genuine random events; namely, irreversible state changes in quantum mechanics. The most well known example is the radioactive decay of atomic nuclei. In such cases one can, using quantum theory, determine the probability that an atom's nucleus will decay within a certain time interval. Furthermore, there is no additional knowledge that could change this probability. If quantum mechanics is a correct and complete theory, then we have a case of maximal

⁶ It is not easy to exactly describe what is meant by a genuinely random event. Even most random events, e.g. computer generation of random numbers, can be shown to be determined by some complicated rule. One definition, due to Chaitin, is to say that a sequence is random if it is not possible to formulate an algorithm that produces the sequence and such that the algorithm is shorter than the sequence itself.

specificity, and the relevant probabilities that can be calculated using this theory are objective. Moreover, the explanatory force does not depend on the probability given.⁷

The conclusion seems to be that in situations where we cite a statistical argument as the explanation of some determined event without possessing all relevant information about the event, then one cannot give a strictly objective interpretation of the statistical explanation. That is to say, the explanation has an unavoidable contextual component, just as in the case of Olga's vote. Explanatory force is thus associated with what one knows prior to receiving the explanation, as well as any other information given when the explanans is presented. But in a case where the statistical law is an expression of genuine randomness and not a measure of our ignorance, one can give an objective and non-contextual analysis of the explanation.

8.7 Action Explanations

Human actions must be treated as a specific category when talking about explanations. This is because actions are intentional phenomena by nature. To explain an action is to view a human as an *agent*, whose actions can be judged in terms of what is rational or irrational. There are certainly exceptions to this rule, such as when one explains the actions of a psychotic person in terms of disturbances in the mechanisms of the brain. Yet we resort to this sort of causal-mechanical explanations only when the usual rational explanations, given in terms of beliefs and desires, fail.

We explain actions by giving the agent's reasons, i.e., citing the relevant motives or goals associated with those actions; this is more or less common sense. If we are unable to give an explanation in terms of motives, and instead attempt to give a biological explanation, we treat the action as behaviour. The difference is that when we describe something as behaviour instead of describing it as an action, we omit intentional components that we include, explicitly or implicitly, when we describe something as an action. Put another way, one can say that the concepts of *action* and *explanation in terms of motives* are logically connected.

We explain an action by citing its purpose or motive, as well as the agent's beliefs, which function as the link between the action and that action's goal. This means that the explanans contains two intentional elements: *goals* and *beliefs*. Thus explanations in terms of motives have the following form:

Agent A wants to achieve goal G

Agent A believes the best way to achieve G is to perform action H

Agent A performs action H

⁷ Peter Railton has observed that the reference class problem can be avoided in cases where probabilities are interpreted as propensities, see Railton (1978). And since transition probabilities in quantum mechanics can be calculated without knowledge about frequencies, they can be viewed as propensities.

There is no guarantee that A in fact performs H, even if she wants to achieve her goal and believes that the best way to reach that goal is by performing H. For example, she could lack the necessary will, or have other goals that conflict with action H. Therefore, this explanation is not a logical inference of the explanandum from the explanans.

Suppose that we observe action H, or rather its behavioural component. In order to explain this action, we must first interpret it as a conscious action by ascribing an intention to perform the act to the agent concerned. We must then find out what beliefs and goals the agent has. However, this is not sufficient. For the explanation to be complete we must know that the action was performed for *just these reasons* and no others. It is not uncommon that the true reasons for an action differ from those explicitly stated, or from those that seem the most reasonable. Thus explanation inevitably contains elements of interpretation as regards the agent's actions and statements.

Notice also that the explanans contains an element of direction towards the future, toward the agent's desired state. Thus some have concluded that that which explains an action lies in the future. However, this view confuses the content of a desire with the mental act of having that desire. The act of having a desire occurs before, or perhaps simultaneously with, the action in question, though the content of this desire is identified by a description of a possible future state of affairs. That which explains an action is the fact that an agent has a certain motive, and this mental state occurs prior to the action. That the content of a motive is directed toward the future doesn't imply that the motive itself is held in the future. So the common argument that action explanations cannot be a species of causal explanations since causes precede their effects and that action explanations are future-oriented, is not valid: having a goal and a belief could cause an action though their content be about future states of affairs.

But there is another reason for not holding action explanations to be causal explanations: causal explanations require causal laws, and there seem to be no strict laws in the mental realm, only rough regularities, such as 'if one is successful in a difficult endeavour, then one becomes satisfied.'

Davidson has in his (1963) defended the common sense view that reasons, i.e., desires and beliefs, explains actions and that it is a species of causal explanation, without reliance of laws. Several philosophers have argued that Davidson is wrong on this point, holding that reasons, i.e., desires and beliefs, are not causes of actions. There has been an intense debate about this issue but no clear conclusion has generally been accepted.

In Chap. 1 I claimed that action explanations are, presumably, the original model for explanations of all phenomena. Natural phenomena were explained by the will of gods, spirits, or other conscious beings in a prime example of overgeneralization. It was a huge step forward when people realized that explanations in terms of motivations were not applicable beyond human actions.

8.8 Pragmatic Explanations

All the theories of explanation hitherto considered take for granted that the explanatory relation is only a relation between explanans and explanandum and that it is possible to determine whether an argument is an explanation without considering the context of the act being explained. Peter Achinstein has formulated this idea in the following way: when we have been convinced that the statements which make up the explanans are true, we should be able to determine, a priori, whether or not the purported explanation really is an explanation.⁸ And Achinstein has argued – convincingly in my view – that for all the well-known models of explanation there are examples where this condition is not met. This indicates that the context of explanation has greater import than normally assumed. This is the basic idea in Bas van Fraassen's account of explanation, which is presented in his book *The Scientific Image*.

Van Fraassen claims that the question as to whether an explanation is correct can be answered only when it is given in the following form: is the given explanation good, or correct, *in this context*? The contextual component is made up of what the questioner knows or believes prior to looking for an explanation. According to van Fraassen, an explanation is *an answer to a why-question*. Thus the theory of explanations is a part of the theory of questions and answers.

Van Fraassen views the following as a typical explanatory situation:

Question Why P?

Answer P is the case, as opposed to Q, R, S, . . . , because A.

P is called the *topic*, which is the state of affairs about which the questioner is inquiring. Furthermore, the questioner assumes a set of possible, although not actualized, alternative states {Q, R, S, . . .}, which together with P are called the *contrast class*. In many cases these alternatives may be implicit. The why-question is thus formulated with the implicit assumption that P is the only actual state of affairs. Hence the first condition for answering the question 'why P?' is thus that 'P' is true and that all other alternatives in the contrast class are false. Suppose someone were to ask you 'why is the sky green?' You could not reasonably explain why the sky is green, and thus the only rational answer would be to reject the implicit claim that the sky is green. The second condition is that the answer, A, must also be true, and the third is that the answer must be *relevant* to the question. In other words, there must exist a *relevance relation* between the question and the answer. What this condition amounts to is not clearly expressed by van Fraassen. Two critics, Kitcher and Salmon, have shown that if one does not place any requirements on the relevance relation, one can construct a relevance relation such that any arbitrary true statement can answer any arbitrary question!⁹ Kitcher and Salmon drew the conclusion that one must impose certain requirements on the relevance relation.

⁸ Achinstein (1981, 1983).

⁹ Kitcher, P., & Salmon, W. (1987). Van Fraassen on explanation. *Journal of Philosophy*, 84, 315–330.

However, there is a question as to whether these critics have correctly understood van Fraassen's intention with his pragmatic model of explanation. Van Fraassen has not replied on this critique, but a defence is easily come by. A consequence of his pragmatic view of explanation is that the relevance relation between question and answer is a pragmatic aspect of the situation; that is, the answer must be deemed relevant *by the questioner in a given situation*. This means that one cannot say, in general, that a certain answer is relevant or irrelevant to a given question, but only that an answer can be relevant to some person in a certain situation while being irrelevant to another person in a different situation. There are therefore no criteria by which we can determine, independent of context, whether or not a proposed explanation is good. This means that it is not possible to place any universally applicable or context-independent requirements on the relevance relation. The question as to whether some explanation is good or not is incomplete. What is meaningful to ask is whether a certain explanation, in a given situation, is acceptable for a particular person. (The only non-contextual requirement on an explanation is that the statements that make up the explanans are true.) According to van Fraassen, there is no such thing as *the correct explanation*. This also implies that a theory's explanatory force is not something that can be used as evidence for the truth or believability of a theory. Explanations are thus uninteresting as regards epistemology, according to van Fraassen, which leads him to reject explanatory force as a relevant factor in comparing two competing theories.

The conclusion is that if van Fraassen is correct, explanations are irrelevant to scientific epistemology; and consequently, we should not devote time and effort to discuss them. According to van Fraassen, the central question in science and the philosophy of science is, 'which theories agree with the empirical data?' Whether a theory has strong or weak explanatory force is a question about what the scientific community considers to be reasonable or obvious. These aspects are not relevant to the question about whether scientific theories agree with empirical data.

Van Fraassen claims that this pragmatic theory solves two important problems affecting most theories of explanation. The first problem is that not all requests for explanation are reasonable. The solution to this problem is easily found in van Fraassen's theory, since a condition for it to be meaningful to answer the question 'why P?', and so reasonable to pose the question in the first instance, is that P must be true.

The other advantage of the pragmatic theory is, according to van Fraassen, that it accounts for the asymmetry of explanation. If A is a relevant answer to the question 'Why P?', then P is not normally a relevant answer to 'why A?'

Arguably, van Fraassen's pragmatic theory is quite successful as regards explanations in ordinary contexts. However, is it not reasonable to say that, in certain scientific situations, the contextual component is either minimal or non-existent? In other words, in a situation where a scientific explanation is requested, one expects an objective, context-independent, and generally acceptable explanation with respect to all relevant scientific theories. For example, one could claim that the explanation of some physical phenomenon should provide the objective physical causes of the phenomenon in question. Van Fraassen could perhaps reply here that

the discussion of causation in the previous chapter showed precisely that the choice of a cause for some event is context-dependent, and that the only context-independent proposition one can make is in regard to what factors make up necessary conditions. In response, one could say that all language use is context-dependent, and thus this is nothing unique to explanation. Furthermore, all of these obvious necessary conditions show up in a given explanation as so-called *Ceteris Paribus clauses*, which will be discussed in Chap. 10.

8.9 Summary

The word ‘explanation’ is used in many different ways in various situations. All together, ‘scientific explanation’ concerns the explanation of phenomena, and not the explanation of the content of different concepts.

There are two fundamental conflicts in the debate over explanation. The first is whether there is a general model that applies to all kinds of scientific explanation. According to Hempel, the D-N model is such a model, which he takes to be generally applicable to all sciences. In response to his critics, who argue that this model does not fit historical explanations, Hempel states that those explanations found in history are not complete; they are merely sketches of explanations, which lack many details. Once these details are filled in, the explanation should fulfil the requirements of the D-N model. The counterargument to this response is the following: if we are to construct a model for scientific explanation, we must first know the actual nature of explanations in each discipline. We cannot simply try to fit explanations into a pre-established model.

The second fundamental dispute concerns whether the concept of scientific explanation contains a contextual component. Hempel, and many of his critics, have taken for granted that the concept of scientific explanation is a non-contextual concept, whereas Achinstein and van Fraassen have claimed that the contextual component is essential. The consequence of this latter view is that explanatory force of a theory cannot be used as an evaluative parameter, since it is relative to the background knowledge of those receiving the explanation. Contrary to this view, one could argue that there is a certain basic perspective shared by all active scientists in a given discipline, which is sufficient for the discussion of a theory’s explanatory force independently of context, within that discipline. In van Fraassen’s terminology, the relevance relation is decided by the theoretical perspective of the active researchers in a given discipline.

Exercises

Below are a number of examples of explanations, some of which are very sketchy. Discuss which model best fits each explanation.

1. Between 1850 and 1920, over one million Swedes immigrated to USA. Why? Historians claim that there were essentially two types of factors that explain this immigration. The first type is the push-factors, such as overpopulation, poverty,

the state church's oppression of 'free-church' movements, etc. The second type was the pull-factors; cultivation of the prairie gave many people the possibility of earning a decent living, leading early immigrants to write home about their success. Once emigration from Sweden was under way, the door was open for those who had family member or friends already settled in the U.S. to travel west and join them.

2. For more than 40 years, Pripps (a Swedish brewing company) had been licensed to produce Coca Cola in Sweden. In 1997, Coca Cola voided Pripps' contract and decided to build their own factory to serve the Scandinavian market. Why did Coca Cola cancel the contract? An explanation given by the press is that Coca Cola was unhappy with its market share in Sweden, since they had a smaller portion of the soft drink market in Sweden than in other comparable countries. This is because Pripps, who dominated the soft drink market at the time, competed with Coca Cola in regards to some similar products. Thus Pripps was not interested in increasing Coca Cola's market share, as this would result in lower sales of Pripps' products.
3. Why did a bloody war in Bosnia follow the fall of Yugoslavia? A not too unreasonable sketch of an explanation is the following. At the fall of communism, Slobodan Milosovic, who was then party chief of the Republic of Serbia, saw where change was headed and that he was in risk of losing power. He chose to appeal to Serbian-nationalist sentiments in order to acquire a new political platform. Milosovic was successful, and the result was that many people dreamed of a Greater Serbia as the successor of Yugoslavia. Since the Serbians practically had control over the Yugoslavian army, their leader chose to start a civil war with the knowledge that they had superior military resources on their side. The goal was to create a cohesive area in Bosnia devoid of Muslims and then join with the remainder of Yugoslavia, thereby building themselves a dream state for all Serbians.
4. Why are carrots orange? Answer: The colour of carrots comes from carotene molecules. All molecules contain electrons that are situated at difference energy levels. In a carotene molecule, the difference in energy between the highest level and the next highest level is about 3 eV. When light hits a carrot, the light ray is absorbed with a frequency that corresponds to the energy gap, according to the formula $E = hf$, where h is Planck's constant and f is the frequency. Blue light has just the right frequency corresponding to the energy gap so that it is absorbed. Thus when we look at a carrot, it looks orange, since the colour complement to the absorbed colour dominates the reflected light.

Further Reading

- Achinstein, P. (1981). Can there be a model for explanation? *Theory and Decision*, 13, 201–227.
 Achinstein, P. (1983). *The nature of explanation*. New York: Oxford University Press.

- Cartwright, N. (1999). *The dappled world. A study of the boundaries of science*. Cambridge: Cambridge University Press.
- Davidson, D. (1980 [1963]). Actions, reasons, and causes. In *Essays on actions and events* (pp. 3–20). New York: Clarendon Press.
- Hempel, C. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.
- Hillel-Ruben, D. (Ed.). (1993). *Explanation. Oxford readings in philosophy*. Oxford: Oxford University Press.
- Pitt, W. (Ed.). (1988). *Theories of explanation*. Oxford: Oxford University Press.
- Railton, P. (1978). A deductive-nomological model of probabilistic explanation. *Philosophy of Science*, 45, 206–226.
- Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
- Salmon, W. (1998). *Causality and explanation*. Oxford: Oxford University Press.
- van Fraassen, B. (1980). *The scientific image*. Oxford: Clarendon Press.