



# Proposing Explanations

## EXPLANATION

**T**he process of proposing and then testing new explanations is at the heart of scientific method. In this chapter we will look at a number of types of explanation with which scientists deal. We will also consider how scientists respond when confronted with rival explanations for a single set of facts. Then, in Chapters 4 and 5 we will look into the way proposed explanations are put to the test. But first we need to do a bit of groundwork by clarifying just what it means to speak of a scientific explanation.

When we ask for an explanation, we could be asking for a number of things. If I'm late for an appointment with you, for example, you might ask me to explain why I'm late. Here, what you want is my excuse. Or, to take another example, you might ask your math teacher to explain how to solve a particularly nasty problem. Here you are asking to be shown how to do something. But suppose I were to bring to your attention the following rather curious fact. In many states, the letter "O" is not used on automobile license plates. You might wonder why this is the case and so you might ask me to explain. In effect, you are asking neither for an excuse nor to be taught how to do something. Instead, you are asking for the reason why something is the case, the reason many states have adopted this somewhat curious policy. (You have probably figured out the explanation for this puzzle. Most license plates are a combination of letters and numbers and the letter "O" could easily be confused with zero.) In speaking of

a “scientific explanation” we are speaking of an explanation in this latter sense: an account of how or why something is the case.

Two terms often associated with explanations in science are *theory* and *hypothesis*. Both theories and hypotheses are explanatory in nature but there are some important differences between the kinds of explanations each is used to describe. “Hypothesis” may be used to refer to anything from a vague hunch to a finely detailed conjecture. In general, however, the point of characterizing a claim as an hypothesis is to note that there is something tentative and unproven about that claim. Thus, if I believe there is intelligent life in the universe somewhere other than earth, my belief may be termed an hypothesis in light of the fact I can produce no hard evidence in its favor. Proposed though untested explanations are sometimes called *explanatory hypotheses*.

“Theory” does not always imply the kind of tentativeness associated with hypotheses. A theory may be a well developed, well confirmed body of explanatory material, as in the big bang theory, the theory of evolution, or the germ theory of disease. But often people say things like, “That’s only a theory,” meaning roughly, “That’s only your opinion of why so-and-so happened.” To make matters worse, many of the things referred to in science as theories are subject to serious question. In astronomy, for example, one highly questionable alternative to the big bang theory is nonetheless referred to as the steady state theory. What typifies theories in science is the breadth and depth of their explanatory capacity. An hypothesis typically will offer an explanation for a limited range of phenomena. Theories tend to be more general structures capable of explaining a much wider variety of phenomena. Moreover, theories will often contain well confirmed rules and principles that reveal underlying explanatory similarities between apparently quite diverse phenomena. Newton’s theory of motion, for example, can explain the behavior of just about anything with a mass, from the tiniest of particles to the stars and planets.

As you can see, “theory” and “hypothesis” are used to cover a lot of ground, and there is no simple and straightforward line of demarcation between the two. The net effect is that when someone speaks of a theory or an hypothesis, we may not be entirely certain what they mean. We can avoid any potential confusion in what follows by speaking simply of explanations. Explanations which share with hypotheses a kind of tentativeness, we can call *novel* or *proposed* explanations or something similar. Explanations which are well established, like some theories, we may simply characterize as *received*, *established*, *generally accepted*, etc.

New scientific explanations do not arise in an intellectual vacuum; they are occasioned by a desire to make sense of something that is not well understood—often anomalies of the sort discussed in Chapter 2. This point may seem so obvious as to hardly bear remarking. Plainly, if we understand something, there is no sense in attempting to provide it with a novel explanation. Yet two closely related points are worth keeping in mind.

First, we must resist the temptation to think of the anomalous as that which is somehow strange and unfamiliar. Now, anomalous phenomena can indeed be strange and unfamiliar, even spectacularly so. Think of some of the cases

discussed in Chapter 2: crop circles, telekinesis, and the strange events of the Bermuda Triangle. Recently astronomers announced the discovery of a pair of collapsed stars that may be composed entirely of something heretofore thought impossible—free quarks. (Quarks are the building blocks of protons and neutrons and are thought to exist only in pairs and triplets.) However, much that needs explaining is considerably less mysterious and unfamiliar. The world about us is filled with phenomena with which we are more than passingly familiar, but which we do not fully understand.

We are all painfully familiar, for example, with many facts about AIDS—about how it is transmitted and what its effects are. Yet a great deal remains to be discovered about the nature of the virus and the way in which it attacks the human immune system. Perhaps nothing is more familiar in our lives than the simple fact that we are creatures capable of thought and feeling. Yet nothing is more puzzling than the way in which neurological processes in the human brain result in mental states like those involved in thought and feeling. What these examples suggest is that both the unusual and the commonplace are ripe for scientific investigation. It has often been remarked that an essential talent of a good scientific researcher is the ability to discern those mundane facts about our daily lives, the investigation of which may yield new and important insights into the ways in which our world works.

Second, the fact that you or I are puzzled by something does not mean that it is genuinely anomalous. Once again, it may seem we are remarking the obvious. Yet as we shall discover when we consider the way in which explanations are tested, it is not uncommon for a person to propose a novel explanation of something because they are unaware that somebody has already adequately explained the phenomenon in question.

The job of explaining rarely comes to an end in science. An explanation tells us something about how or why a thing happens, but rarely will an explanation be so complete as to leave no further unanswered “whys” or “hows” about the thing in question. To see this, consider the following causal explanation. We all know that the tides are caused in part by the gravitational attraction of the moon. Thus, we can explain the tides by reference to the fact that there is a large amount of water on the surface of the earth, that the earth rotates on its axis, and that the source of gravitational attraction, the moon, moves in orbit around the earth. Now, though our explanation clearly gives us a sense of why there should be two high and low tides roughly every twenty-four hours, it leaves a lot unexplained. What is the process by which gravitating bodies, in this case the moon and the oceans, interact? Put another way, how is it that massive objects such as these have an effect on one another? We might cite here something called the law of gravity: objects tend to attract one another in direct proportion to their masses and in inverse proportion to the square of the distance between them. In a sense, this adds a bit of detail to our explanation. But why should this “law” hold? Why should objects attract one another at all, let alone in just this regular, law-like fashion? Unfortunately, we must leave these questions unanswered, for little is yet known about what physicists today call the “carrier” of gravitational interaction, “the graviton.”

As our last example suggests, explaining one thing in science often leads naturally to the need for new, more fundamental explanations. The moral of this last point is that in science, at any rate, progress is largely a matter of providing better and better approximations of what is going on in nature. Rarely are explanations final or complete in the sense of leaving no additional unanswered questions about what is really going on.

Scientific progress is not always a matter of supplementing received explanations with more subtle but complementary new explanations. The history of science is fraught with instances in which received explanations have been supplanted by novel and radically different ones. One of the most well known examples of the replacement of one explanation with another is the gradual shift from the Ptolemaic conception of the universe to the Copernican.

In the Ptolemaic view, systematized about 140 A.D. by Ptolemy Claudius of Alexandria, the stationary earth stands at the center of the universe and all heavenly objects revolve around the earth. The Ptolemaic view had considerable explanatory power in that by a series of complicated calculations, the motions of all celestial objects known at the time—the sun, the moon, the five innermost planets, and the stars—could be explained, though in ways very different than we would explain them today. For example, careful observation revealed that Mars generally moves eastward across the night sky but occasionally appears to move backward for a bit before resuming its eastward course. In the Ptolemaic view, all celestial objects trace out circular orbits around the earth. Ptolemy explained the backwards, or retrograde, motion of Mars by introducing the notion of an epicycle—a small circular loop in the orbit of Mars such that, from an earthly perspective, Mars would actually appear to stop and then move backwards during its epicycle. A tribute to its explanatory value is the fact that the Ptolemaic view dominated Western thought for more than a thousand years.

In the sixteenth century, however, Nicholas Copernicus, a Polish scientist and astronomer, proposed a new and radically different view of the cosmos. In Copernicus's view, many of the basic assumptions of Ptolemy were wrong. The sun, not the earth, is at the center of things; two of the planets, Mercury and Venus, occupy orbits nearer the sun than does the earth; and, what is more, many celestial motions are to be explained by the fact that the earth rotates on its axis. One advantage of the Copernican view is that it suggests a very different explanation for retrograde motion than does that of Ptolemy. If, as Copernicus suggested, the orbit of Mars is outside that of the earth, then the double motion of Mars with respect to the earth explains the apparent backward motion of Mars. For in Copernicus's view, we observe the motion of Mars from a location that is itself moving through space with the net effect that Mars will on occasion appear to be moving backwards.

There are a number of interesting facts about this particular episode in the history of science. The first, of course, is the enormous shift in thinking about the nature of celestial motions occasioned by the work of Copernicus. One might think the Copernican "revolution," as it is sometimes called, would have ushered in a new level of accuracy and simplicity in the calculation of planetary

motions. But as it turned out, Copernicus's explanation was neither more accurate nor even much simpler than that of Ptolemy. Both views explained roughly the same collection of data about planetary motion. Moreover, like Ptolemy, Copernicus had to introduce a number of epicycles into his work to make his explanation fit the facts. The real value, then, of Copernicus's achievement resides in the simple but profoundly new way of thinking about celestial motion it introduced.

But our story does not end here. Though in rough outline the Copernican view of the universe finally replaced that of Ptolemy, many of the details of the Copernican view were themselves eventually rejected. Copernicus, like Ptolemy, for example, believed that the planets trace out circular orbits around the sun. (In fact, it was this conviction that necessitated the introduction of the occasional epicycle in his calculations.) It remained for Johannes Kepler, nearly a century later, to discover that the planets trace out elliptical orbits around the sun. Kepler thereby reduced the kinds of motion required to explain the observed positions of the planets and did away, finally, with the infamous epicycle. In defense of Copernicus, it must be noted that Kepler had available much more accurate measurements of the movement of the planets than anything available to either Copernicus or Ptolemy. Yet despite the enormous import of Kepler's contributions to our understanding of celestial motion, it remained for astronomers long after the time of Kepler to refine the Copernican world view even further by removing the sun from its exalted position at the center of the universe.

## CAUSES

When we think about what is involved in giving an explanation, the notions of cause and effect come immediately to mind. Indeed, the most obvious kind of explanation is causal explanation. Why, for example, when we were small children, did teeth, carefully tucked under our pillows, vanish only to be replaced by money? Because while we were sleeping our parents removed the teeth and replaced them with money. Why is there a circular crater several miles in diameter in the Arizona desert? Because a large meteor survived intact its trip through the earth's atmosphere; its crash produced the crater. Why is smoking on the increase among young adults? In part, because the tobacco industry targets this segment of the population in much of its advertising. In each of these cases a cause for a particular effect is identified and with each we understand something of why the phenomena in question is the case.

Causal explanations are common in our daily lives. Imagine I've arrived late for a lunch engagement. "Sorry I'm late. The traffic was horrendous," I say. What I am doing here is claiming that something out of my control caused me to be late. Or suppose the street out front of the restaurant where we are meeting is flooded. You venture the guess that all of the drains are clogged with leaves. Your guess involves a causal explanation. The leaves covering the drains have caused the street to flood.

Causal relationships are not always as simple or straightforward as our last example. For one thing, effects can have more than a single cause. It may be, for example, that my lateness was in part caused by a traffic jam. But suppose that while hung up in traffic I ran low on gas and so had to stop and fill up. Suppose also that neither event alone would have made me late. In the jargon of causal research, multiple related causes are referred to simply as “causal factors.” Moreover, effects need not invariably be associated with a given causal factor. We know, for example, that cigarette smoking causes lung cancer despite the fact that some cigarette smokers will not contract lung cancer and that some who will contract lung cancer will not be smokers. As this last example suggests, causal explanations are often about groups, not individuals. The claim that smoking causes lung cancer means that, among people in general (and several kinds of laboratory animals), smoking is one factor that contributes to lung cancer. Finally, causes can be either *remote* or *proximate*. If, say, A causes B, which in turn causes C, A is often referred to as a proximate cause of B and a remote cause of C. So, for example, if I trip and bump into the table where you are seated causing your water glass to spill in your lap, my tripping is a proximate cause of the movement of the table and a remote cause of the mess in your lap.

## CORRELATION

Closely related to the notion of a causal explanation is that of a correlation. Indeed, people often assume that if two things are correlated they are causally linked. But this assumption is often wrong. A correlation is nothing more than a comparison between a pair of characteristics within a population. Those characteristics are correlated if they display some regular, measurable variance. The simplest sort of correlation involves the comparison of two groups, one having a given characteristic and the other lacking it. If a second characteristic occurs at different frequencies in the two groups, it is correlated with one of the two. Suppose, for example, that we compared two groups of people, all between ages 30 and 49. Each member of the first group has completed at least four years of college, while those in the second group have completed less than four. Suppose also that we were able to look at the average annual income of the two groups and were to find that the income of the first group is, on average, 20 percent higher than that of the second group. This means there is a correlation between education and income in the groups of people we have considered.

Correlations can be positive or negative. If a characteristic occurs at a greater frequency in one group than in the other, it is positively correlated with the first group; if it occurs at a lesser frequency, the correlation is negative. By contrast, if the characteristic occurs at roughly the same frequency in both groups, there is no correlation between the characteristic and either group. In our example, we have uncovered a positive correlation between education and income. Suppose instead we had found that the income of those having four or more years of college was actually lower than that of people with less

education. This finding would suggest a negative correlation between the two factors. Had we found no real difference in levels of income, we would have to conclude that, insofar as we can tell, there is no correlation between level of education and income. (This does not mean there is no such correlation. All we can conclude is that our quick check of the data available shows no correlation!)

Correlations can also hold between pairs of characteristics within a single group. Within a group, if two measurable characteristics vary in a somewhat regular and predictable fashion, they are correlated. Suppose, for example, we had at our disposal a large amount of information about the freshman class at a small local college. Examining the data we find what appears to be an interesting relationship between first semester grade point averages (GPA) and scholastic aptitude test (SAT) scores. About 100 students completed the first semester. In most cases, say 75 or so, we find that GPA varies directly with SAT score. That is, if we arrange these 75 students in order of ascending SAT score, we find a corresponding increase in GPAs; the higher the SAT score, the higher the GPA. For the other 25 or so students, we find no regular variance. Some students with relatively high SAT scores have relatively low GPAs and vice versa. Some with average SAT scores have relatively high, some relatively low GPAs. Despite these exceptions, our findings suggest a positive correlation between SAT score and GPA, at least in the group we have examined. Had we found just the reverse—had we found that for most students, GPA diminished when SAT scores increased, we would have uncovered a negative correlation between SAT score and GPA. Suppose instead we were to discover no regular variance between SAT scores and GPAs; many students with relatively high SAT scores had average or low GPAs while many with relatively low SAT scores had average or high GPAs. This would suggest that no correlation exists between SAT score and GPA in the freshman class of the college.

As our last example suggests, correlation is seldom an all or nothing matter. A perfect correlation between two characteristics would require a one-to-one correspondence between changes in the two. (In our example, increases in SAT score would need to be accompanied by increases in GPA in all 100 cases to establish a perfect correlation.) But particularly when groups of subjects are large, the fact that a correlation is somewhat less than perfect does not undercut its potential significance, perhaps as a predictor of one characteristic in cases where we know something about the quantity of the other. Presuming, in our example, that we have uncovered a fairly consistent positive correlation between SAT score and first semester GPA, we may be able to predict something about a new college student's chances of success based on his or her SAT score. But here we need to introduce a crucial note of caution. Any inference we draw about an individual, based on the evidence of a correlation, assumes a causal connection between the correlated characteristics. And this assumption is not always warranted. The fact that two things are correlated does not, by itself, indicate that the two are causally linked.

Why this is so can be seen in the following examples. If we were to examine a group of similar people, say, members of a single trade or profession, we could probably unearth a number of correlations. We might find, for example,