A Beginner's Guide to Scientific Method

FOURTH EDITION

STEPHEN S. CAREY
Portland Community College

WADSWORTH
CENGAGE Learning

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

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.

CAUSATION

One way to explain how or why something has occurred is to give an account of the events leading up to it. 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 the diameter in the Arizona desert? Because a large meteor survived its trip through Earth's atmosphere intact; 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 phenomenon 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. My excuse is to the effect 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 simple or straightforward. Consider some of the complexities we must face in thinking about causes and their effects.

First, effects can be the result of a combination of causes. 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. My being late has been caused by a combination of events.

Second, both causes and effects can be about groups rather than individual facts or events. To claim, for example, that cigarette smoking causes lung cancer is to claim that lung cancer will occur with greater frequency among the group of people who smoke than among those who do not.

Third, effects may result from several distinct causes. We know that cigarette smoking causes lung cancer. But other things—exposure to asbestos and genetic abnormalities, for example—can cause lung cancer as well. In some cases, a series of discrete causes will be responsible for an overall effect, though each factor will be responsible for only part of the effect. In the early 1990s, violent crime rates in the United States fell precipitously after 15 years of constant increase. As it turns out, several factors were responsible for the decline, though each was responsible for only part of the overall decline. Among the more pronounced causal factors were: increased use of capital punishment, increased rates of imprisonment, increased number of police, changes in the illegal drug market, and the legalization of abortion approximately 20 years earlier.[1] A rough estimate is that the hiring of additional police accounted for roughly 10% of the 1990s crime drop and that decreased demand for crack cocaine was responsible for another 15%.

Fourth, effects need not invariably be associated with a given causal factor. Though cigarette smoking is indeed a cause of lung cancer, many cigarette smokers will neither contract nor die of lung cancer. There is today a good deal of evidence that children who drink fluoridated water will have fewer problems with tooth decay than children who do not drink fluoridated water. Though there is clearly a causal link between fluoride and tooth decay, it does not follow that children who drink fluoridated water will be completely free of decay problems.

Fifth, causal explanations can be negative, as in our last example. Fluoridation of the water supply prevents tooth decay. Similarly, many people believe there is a causal link between vitamin C and the common cold. Regular doses of vitamin C, it is claimed, will decrease your chances of contracting a cold.

Finally, causal explanations can involve a sequence of linked events. 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. B in turn is a proximate cause of C. So for example, if I trip and bump into the table where you are seated, causing your water glass to spill into your lap, my tripping is the proximate cause of the movement of the table and the 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. The first group all have completed at least four years of college, while the second have less than four years. 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% 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. Suppos instead we had found that the income of those having four or more years (college was actually lower than that of people with less education. This findin would suggest a negative correlation between the two factors. Had we found n real difference in levels of income, we would have had to conclude that, insof as we can tell, there is no correlation between level of education and incom (This does not mean there is no such correlation. All we can conclude is that o quick check of the data available shows no correlation!)

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

As our last examples suggest, correlation is seldom an all-or-nothing matte A perfect correlation between two characteristics would require a one-to-on correspondence between changes in the two. (In our example, increases in SA score would need to be accompanied by increases in GPA in all 100 cases t establish a perfect correlation.) But particularly when groups of subjects ar large, the fact that a correlation is somewhat less than perfect does not underc its potential significance, perhaps as a predictor of one characteristic in cas where we know something about the breakdown of the other. Presuming, i our example, that we have uncovered a fairly consistent positive correlatio between SAT score and first semester GPA, we may be able to predict somethin about a new college student's chances of success, based on his or her SAT scor But here we need to introduce a crucial note of caution. Any inference we dra about an individual, based on the evidence of a correlation, assumes a caus connection between the correlated characteristics. And this assumption is n always warranted. The fact that two things are correlated does not, by itse 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, a correlation between age and income, established either by showing a regular variance between age and income for the whole group, or by showing that people above and below a certain age have, on average, different income levels. We would probably also find a correlation between age and the use of reading glasses. Given these correlations, it is likely we will also find a correlation between income and the use of reading glasses! Now, none of these correlations seems to be a coincidence. There seems to be a clear link between age and the need for reading glasses. But the link in the other two cases is much more tenuous. Advancing age does not cause one's income level to rise, nor does income have any bearing on the need for reading glasses. The link in these two cases is undoubtedly explained by some other factor or series of factors. For example, in most trades or professions, the longer one works at a job, the more one generally makes. This, then, accounts for the correlation between age and income.

To make matters worse, a correlation may be evidence of nothing more than coincidence, a "mere correlation." This is because unrelated things can vary in regular, measurable ways. For a number of years now, two things have regularly increased: the sale of Burger King Whoppers and the number of minutes per day that children watch television. Come to think of it, recent increases in Whopper sales are correlated, negatively, with a gradual but regular decrease, in the same period, in the number of people who go bowling! And since we are an aging population, I suspect we could also dredge up a correlation between Whopper sales and the purchase of reading glasses. These new correlations, of course, suggest nothing more than the fact that lots of things, many of them not causally related, vary over time in somewhat regular ways.

All of this is not to say that the search for correlations is not an important component of causal research. Indeed, if two things are causally linked, they will be correlated, and so evidence of a correlation may provide some initial evidence for a causal link. But the simple fact that two things are correlated is, by itself, not evidence of a causal link. In Chapter 5 we will look closely at the ways in which claims about causal links and their attendant correlations are tested. For now, it is enough to keep in mind that correlations do not necessarily indicate causal links and, for this reason, are of less explanatory value than are facts about causal links.

Though causal explanation plays a central role in science, there are several other ways of explaining, all of which have a distinct role to play in scientific investigation. Consider another pair of examples. Why do our eyelids blink open and shut several times every minute? To keep the surface of the eye moist. Why does a gun "kick" as it is discharged? Because of a well established physical law: for every action there is an equal and opposite reaction. Neither of these explanations involves a cause, at least in any straightforward sense. We normally think of causes as events that precede the things they bring about. Physical laws do not cause things to happen in this sense. Nor is the blinking of an eyelid caused by the fact that it thereby keeps the eye moist.

An explanation may focus on the function or role a thing plays in some larger enterprise or on a law that accounts for the behavior in question. These

QUICK REVIEW 3.1 Causation and Correlation

Causation

Two things are causally linked if one proceeds and is responsible for the other. Suppose your car won't start because its battery is dead. The dead battery is the cause and your car's failure to start, the effect. Effects can have more than a single cause, and there may be many causes for similar effects. Several causal factors are responsible for the behavior of the stock market and a market decline can be caused by a variety of factors. Causal relationships can hold between individual events or between large classes of events as in the claim that megadoses of vitamin C can reduce occurrences of the common cold. If events are causally linked they will be correlated, but correlations do not necessarily indicate causal links.

Positive Correlation

In two populations, P and Q are positively correlated if a greater percentage of Ps than non-Ps have Q. Suppose that nationwide, people with cell phones have, on average, a higher income than people without cell phones. Cell phone ownership and income are positively correlated.

In a single population, if a regular increase in one trait, P, is accompanied by a regular increase in another, Q, then the two are positively correlated. Suppose worker productivity at a plant increases as pay increases, though with some exceptions. Worker productivity is positively correlated with income.

Negative Correlation

In the two populations, P and Q are negatively correlated if a smaller percentage of Ps than non-Ps have Q. Suppose regular users of the local public library (once or more a month) watch, on average, much less TV than sporadic or nonusers of the library. TV-watching and library use are negatively correlated for the group in question. In a single population, P and Q are negatively correlated if a regular increase in P is accompanied by a regular decrease in Q. Suppose that the number of visits to the library per month increases as the average number of hours watching TV decreases. Library use and TV watching are negatively correlated.

No Correlation

In two populations, P and Q are not correlated if there is no difference in levels of Q in P and not-P. If equal percentages of males and females are left-handed, there is no correlation between left-handedness and gender. In a single population, two traits are not correlated if there is no regular variance between the occurrence of the two. Suppose we were to record both the number of checks written and the number of soft drinks consumed per month by a randomly chosen group of people. We would probably find no evidence that variation in one trait is a predictor of a variation in the other. This suggests there is no correlation between the two.

Perfect Correlation

An invariant relation between two traits; for every change in one trait there is a consistent change in the other. In most species of trees, age in years is perfectly correlated with the number of rings in the tree's trunk; the older the tree the greater the number of rings, without exception.

strategies were used in the examples just above. Explanations can also make reference to underlying processes and causal mechanisms, techniques often used to enrich causal explanations. Since each of these ways of explaining plays an important role in scientific inquiry, let's take a closer look at what each involves.

CAUSAL MECHANISMS

Earlier we noted that a causal explanation can involve a series of linked events. One cause can be more remotely connected to its effect than another, more proximate cause. A causal mechanism is nothing more than a series of proximate causes that intervene between a remote cause and its effect. Researchers often comment that though they have evidence for a causal link, they lack a clear understanding of the mechanism involved. This tells us that the sequence of causes which have led to the effect has not been fully fleshed in. Cigarette smoking, for example, is known to be linked to lung cancer. Yet despite the fact that we are quite confident there is a link between smoking and lung cancer, little is know about the mechanism—the physiological process—by which the carcinogens ("carcinogen" just means "cancer-causing agent") in cigarette smoke lead to uncontrolled cell growth in the lungs of the smoker.

A recent study revealed an apparent causal connection between aspirin consumption and the risk of heart attack. According to the study, men who take a single buffered aspirin every other day have a 50% lower chance of having a heart attack than do men who do not take aspirin. Here the connection between aspirin consumption and risk of heart attack seems to be fairly well documented. As it turns out, the causal mechanism by which aspirin reduces the risk of heart attack is also well understood. Aspirin interferes with the first stage of the blood's clotting process. Now, many heart attacks are caused by blood clots in damaged arteries. It seems that when the thin inner wall of an artery is damaged, aspirin inhibits the tendency of minute blood platelets to clot over the damaged area. Thus, aspirin reduces the clotting effect that can lead to serious heart attack.

To take a very different example, one more closely related to every day life, imagine the following.[2] A friend applied for a job she really wanted to get. Yet now she tells us she finds the job utterly uninteresting and probably wouldn't accept it even if it were offered to her. Why the change in attitude? We discover subsequently that our friend learned she had no chance of getting the job. But how, if at all, did this bring about her change in attitude about the job? The answer may well lie in a causal mechanism, often called cognitive dissonance reduction, that makes people cease desiring that which they cannot get; you may be familiar with this mechanism under its more common name, sour grapes. Having learned she wouldn't get the job, our friend adjusted her desires thereby reducing the dissonance caused by wanting something she could not have. No doubt, the notion of cognitive dissonance reduction is a bit less precise than is the mechanism invoked to explain the connection between aspirin and heart disease, and for that reason it would be more difficult to test. But such

psychological mechanisms nonetheless play an important role in our attempts at explaining why people behave as they do.