THE FUNDAMENTALS OF

# Political Science Research

*Third Edition*

**Paul M. Kellstedt**

Texas A&M University

**Guy D. Whitten**

Texas A&M University

CAMBRIDGE
UNIVERSITY PRESS

# 3 Evaluating Causal Relationships

## OVERVIEW

Modern political science fundamentally revolves around establishing whether there are *causal relationships* between important concepts. This is rarely straightforward, and serves as the basis for almost all scientific controversies. How do we know, for example, if economic development causes democratization, or if democratization causes economic development, or both, or neither? To speak more generally, if we wish to evaluate whether or not some $X$ causes some $Y$, we need to cross four causal hurdles: (1) Is there a credible causal mechanism that connects $X$ to $Y$? (2) Can we eliminate the possibility that $Y$ causes $X$? (3) Is there covariation between $X$ and $Y$? (4) Have we controlled for all confounding variables $Z$ that might make the association between $X$ and $Y$ spurious? Many people, especially those in the media, make the mistake that crossing just the third causal hurdle – observing that $X$ and $Y$ covary – is tantamount to crossing all four. In short, finding a relationship is not the same as finding a *causal* relationship, and causality is what we care about as political scientists.

*I would rather discover one causal law than be King of Persia.*
—Democritus (quoted in Pearl, 2000)

## 3.1 CAUSALITY AND EVERYDAY LANGUAGE

Like that of most sciences, the discipline of political science fundamentally revolves around evaluating causal claims. Our theories – which may be right or may be wrong – typically specify that some independent variable causes some dependent variable. We then endeavor to find appropriate empirical evidence to evaluate the degree to which this theory is or is not supported. But how do we go about evaluating causal claims? In this chapter and the next, we discuss some principles for doing this. We focus on the logic of causality and on several criteria for establishing with some confidence the degree to which a causal connection exists between two variables. Then, in Chapter 4, we discuss various ways to design research that help us to investigate causal claims. As we pursue answers to questions about causal relationships, keep our "rules of the road" from Chapter 1 in your mind, in particular the admonition to consider only empirical evidence along the way.

It is important to recognize a distinction between the nature of most scientific theories and the way the world seems to be ordered. Most of our theories are limited to descriptions of relationships between a *single* cause (the independent variable) and a *single* effect (the dependent variable). Such theories, in this sense, are very simplistic representations of reality, and necessarily so. In fact, as we noted at the end of Chapter 1, theories of this sort are laudable in one respect: They are parsimonious, the equivalent of bite-sized, digestible pieces of information. We emphasize that the great majority of our theories about social and political phenomena are **bivariate** – that is, involving just two variables.

But social reality is *not* bivariate; it is **multivariate**, in the sense that any interesting dependent variable is caused by more than one factor. ("Multivariate" simply means "many variables," by which we mean involving more than two variables.) So although our theories describe the proposed relationship between some cause and some effect, we always have to keep in the forefront of our minds that the phenomenon we are trying to explain surely has many other possible causes. And when it comes time to design research to test our theoretical ideas – which is the topic of Chapter 4 – we have to try to account for, or "control for," those other causes. If we don't, then our causal inferences about whether our pet theory is right – whether $X$ causes $Y$ – may very well be wrong.[1] In this chapter we lay out some practical principles for evaluating whether or not, indeed, some $X$ does cause $Y$. You also can apply these criteria when evaluating the causal claims made by others – be they a journalist, a candidate for office, a political scientist, a fellow classmate, a friend, or just about anyone else.

Nearly everyone, nearly every day, uses the language of causality – some of the time formally, but far more often in a very informal manner. Whenever we speak of how some event changes the course of subsequent events, we invoke causal reasoning. Even the word "because" implies that a

[1] Throughout this book, in the text as well as in the figures, we will use arrows as a shorthand for "causality." For example, the text "$X \rightarrow Y$" should be read as "$X$ causes $Y$." Oftentimes, especially in figures, these arrows will have question marks over them, indicating that the existence of a causal connection between the concepts is uncertain.

causal process is in operation.[2] Yet, despite the ubiquitous use of the words "because," "affects," "impacts," "causes," and "causality," the meanings of these words are not exactly clear. Philosophers of science have long had vigorous debates over competing formulations of "causality."[3]

Although our goal here is not to wade too deeply into these debates, there is one feature of the discussions about causality that deserves brief mention. Most of the philosophy of science debates originate from the world of the physical sciences. The notions of causality that come to mind in these disciplines mostly involve **deterministic relationships** – that is, relationships such that if some cause occurs, then the effect will occur *with certainty*. In contrast, though, the world of human interactions consists of **probabilistic relationships** – such that increases in $X$ are associated with increases (or decreases) in the probability of $Y$ occurring, but those probabilities are not certainties. Whereas physical laws like Newton's laws of motion are deterministic – think of the law of gravity here – the social sciences (including political science) more closely resemble probabilistic causation like that in Darwin's theory of natural selection, in which random mutations make an organism more or less fit to survive and reproduce.[4]

What does it mean to say that, in political science, our conceptions of causality must be probabilistic in nature? When we theorize, for example, that an individual's level of wealth causes her opinions on optimal tax policy, we certainly do not mean that *every* wealthy person will want lower taxes, and *every* poor person will prefer higher taxes. Consider what would happen if we found a single rich person who favors high taxes or a single poor person who favors low taxes. (Perhaps you are, or know, such a person.) One case alone does not decrease our confidence in the theory, let alone disprove it entirely. In this sense, the relationship is probabilistic, not deterministic. Instead of saying deterministically that "wealthy people will prefer lower taxes, and poorer people will prefer higher taxes," we say, probabilistically, that "wealthy people are more likely to prefer lower taxes, whereas poorer individuals are more likely to prefer higher taxes."

2 This use of terms was brought to our attention by Brady (2002).

3 You can find an excellent account of the vigor of these debates in a 2003 book by David Edmonds and John Eidinow titled *Wittgenstein's Poker: The Story of a Ten-Minute Argument Between Two Great Philosophers.*

4 Nevertheless, in reviewing three prominent attempts within the philosophy of science to elaborate on the probabilistic nature of causality, the philosopher Wesley Salmon (1993, p. 137) notes that "In the vast philosophical literature on causality [probabilistic notions of causality] are largely ignored." We borrow the helpful comparison of probabilistic social science to Darwinian natural selection from Brady (2004).

**YOUR TURN: Deterministic or probabilistic?**

Before the 2012 presidential election, many observers noted that no US president since World War II had been reelected with an unemployment rate above 8 percent.

Identify the causal claim embedded in this statement. Is it deterministic or probabilistic? Is that a problem?

Take another example: Scholars of international conflict have noticed that there is a statistical relationship between the type of regime a country has and the likelihood of that country going to war. To be more precise, in a series of studies widely referred to as the "democratic peace" literature, many researchers have noticed that wars are much less likely to break out between two regimes that are democracies than between pairs of countries where at least one is a nondemocracy. To be perfectly clear, the literature does not suggest that democracies do not engage in warfare at all, but that democracies don't fight other democracies. A variety of mechanisms has been suggested to explain this correlation, but the point here is that, if two democracies start a war with one another next year, it would be a mistake to discard the theory. A deterministic theory would say that "democracies don't go to war with one another," but a more sensible probabilistic theory would say that "democracies are highly unlikely to go to war with one another."

In political science there will always be exceptions, because human beings are not deterministic robots whose behaviors always conform to lawlike statements. In other sciences in which the subjects of study do not have free will, it may make more sense to speak of laws that describe behavior. Consider the study of planetary orbits, in which scientists can precisely predict the movement of celestial bodies hundreds of years in advance. The political world, in contrast, is extremely difficult to predict. As a result, most of the time we are happy to be able to make statements about probabilistic causal relationships.

Indeed, approaches to studying causal relationships are still being refined today. For example, the statistician Donald Rubin (1974) has developed a rigorous framework for evaluating what are called "the effects of a cause." It is based on the understanding that a causal effect can be measured by examining the different potential outcomes for a case, depending on the assignment condition that a case receives. In an ideal setting, if we want to know if $X$ causes $Y$, we would like to observe outcomes ($Y$) for the same cases with all values of the treatment ($X$).[5] The main problem with causal inference is that we cannot observe multiple

[5] Rubin would call the independent variable the "treatment."

outcomes for the same case. What must be done, then, is to formulate methods to facilitate comparisons between groups so that the assignment between groups does not affect our conclusions about the relationship between $X$ and $Y$. More on this in Chapter 4.

What all of this boils down to is that the entire notion of what it means for something "to cause" something else is far from a settled matter. In the face of this, should social scientists abandon the search for causal connections? Not at all. What it means is that we should proceed cautiously and with an open mind, rather than in some exceedingly rigid fashion.

## 3.2 FOUR HURDLES ALONG THE ROUTE TO ESTABLISHING CAUSAL RELATIONSHIPS

If we wish to investigate whether some independent variable, which we will call $X$, "causes" some dependent variable, which we will call $Y$, what procedures must we follow before we can express our degree of confidence that a causal relationship does or does not exist? Finding some sort of covariation (or, equivalently, correlation) between $X$ and $Y$ is not sufficient for such a conclusion.

We encourage you to bear in mind that establishing causal relationships between variables is not at all akin to hunting for DNA evidence like some episode from a television crime drama. Social reality does not (often) lend itself to such simple, cut-and-dried answers. In light of the preceding discussion about the nature of causality itself, consider what follows to be guidelines as to what constitutes "best practice" in political science. With any theory about a causal relationship between $X$ and $Y$, we should carefully consider the answers to the following four questions:

1. Is there a credible causal mechanism that connects $X$ to $Y$?
2. Can we rule out the possibility that $Y$ could cause $X$?
3. Is there covariation between $X$ and $Y$?
4. Have we controlled for all **confounding variables** $Z$ that might make the association between $X$ and $Y$ **spurious**?[6]

Let's discuss these in turn.

First, we must consider whether it is believable to claim that $X$ *could* cause $Y$. In effect, this hurdle represents an effort to answer the "how" and "why" questions about causal relationships. To do this, we need to go through a thought exercise in which we evaluate the mechanics of how

[6] A "confounding variable" is simply a variable that is correlated with both the independent and dependent variables and that somehow alters the relationship between those two variables. "Spurious" means "not what it appears to be" or "false."

$X$ would cause $Y$, or how varying the levels of $X$ might cause the levels of $Y$ to vary. What is the process or mechanism that, logically speaking, suggests that $X$ might be a cause of $Y$? In other words, what is it specifically about having more (or less) of $X$ that will in all probability lead to more (or less) of $Y$? The more outlandish these mechanics would have to be, the less confident we are that our theory has cleared this first hurdle. Failure to clear this first hurdle is a very serious matter; the result being that either our theory needs to be thrown out altogether, or we need to revise it after some careful rethinking of the underlying mechanisms through which it works.

What do we mean by a "credible causal mechanism"? Perhaps two examples will help shed some light on what is an admittedly cumbersome phrase. In our example from Chapter 1 on the theory of economic voting, which attempts to connect variations in economic performance ($X$) to an incumbent party's reelection vote percentage ($Y$), can we identify answers to this question? "How, specifically, might varying economic conditions cause an incumbent's vote shares to vary?" Yes, we can. If the population of a country values an economy that is performing well – with strong growth, low inflation, and low unemployment, for example – and if the voters hold the governing party, at least in part, responsible for the management of the economy, then voters might base their votes for or against the incumbent party on how well or how poorly the economy is doing. If the economy does well, more voters might reward the incumbent party for competent management of the economy with their votes; if the economy fares poorly, more voters might punish the incumbent party for inept management of the economy, and vote for the opposition. This series of statements would qualify as a "credible causal mechanism" in this case, and we would clear causal hurdle 1. However, just because something is "credible" – that is, believable – doesn't necessarily make it true, or show that the theory is right. It just means that we've identified a plausible, potentially true, mechanism of how $X$ might cause $Y$.

To further illustrate the point, let's consider a different example. Ice cream consumption varies over the course of the year, as you might expect; most of us eat more ice cream in the hotter summer months, and less ice cream during the colder winter. Homicides, in many large cities, follow a similar pattern, with more murders in the summer months and fewer in the winter. What if we wanted to investigate the possibility that the monthly variation in ice cream consumption ($X$) *caused* the variation in homicides ($Y$)? As with any question about causality, our first step should be to ask if we can identify a credible causal mechanism that might connect changes in ice cream consumption with shifts in the murder rate. And that's really difficult, if not impossible, to do in this case. It's true, we acknowledge, that we could get a bit cheeky and say that, after all, when people eat more ice cream, they might get fueled rage from all of the high-fructose corn syrup, and therefore be more likely to commit murders. But even though that might make you chuckle, that's just not believable in any serious way. It's not something you can say with a straight face. So, in this case, our "theory" – note the scare quotes here, because we don't really have a proper theory – would not cross the first causal hurdle. We'd like to emphasize that it is only worth proceeding to the second question once we have a "yes" answer to this first question. That is, if you cannot specify a believable, potentially true process by which varying *X* might cause *Y* to vary, then stop now, and work on formulating a theoretical explanation about the relationship.

Second, and perhaps with greater difficulty, we must ask whether we can rule out the possibility that *Y* might cause *X*. As you will learn from the discussion of the various strategies for assessing causal connections in Chapter 4, this poses thorny problems for some forms of social science research, but is less problematic for others. Occasionally, this causal hurdle can be crossed logically. For example, when considering whether a person's gender (*X*) causes him or her to have particular attitudes about abortion policy (*Y*), it is a rock-solid certainty that the reverse-causal scenario can be dismissed: A person's attitudes about abortion does not "cause" them to be male or female. If our theory does not clear this particular hurdle, the race is not lost. Under these circumstances, we should proceed to the next question, while keeping in mind the possibility that our causal arrow might be reversed.

Throughout our consideration of the first two causal hurdles, we were concerned with only two variables, *X* and *Y*. The third causal hurdle can involve a third variable *Z*, and the fourth hurdle always does. Often it is the case that there are several *Z* variables.

For the third causal hurdle, we must consider whether *X* and *Y* covary (or, equivalently, whether they are correlated or associated). Generally speaking, for *X* to cause *Y*, there must be some form of measurable association between *X* and *Y*, such as "more of *X* is associated with more of *Y*," or "more of *X* is associated with less of *Y*." Demonstrating a simple bivariate connection between two variables is a relatively straightforward matter, and we will cover it in Chapters 8 and 9. Of course, you may be familiar with the dictum "Correlation does not prove causality," and we wholeheartedly agree. It is worth noting, though, that bivariate correlation is normally an essential component of a causal relationship.

But be careful. If you read the above paragraph carefully, you'll have noticed that we said that correlation is "normally" – not "universally" – a component of a causal relationship. It is possible for a causal relationship to exist between *X* and *Y* even if there is no bivariate association between

*X* and *Y*. Thus, even if we fail to clear this hurdle, we should not throw out our causal claim entirely. Instead, we should consider the possibility that there exists some confounding variable *Z* that we need to "control for" before we see a relationship between *X* and *Y*. Whether or not we find a bivariate relationship between *X* and *Y*, we should proceed to our fourth and final hurdle.

Fourth, in establishing causal connections between *X* and *Y*, we must face up to the reality that, as we noted at the outset of this chapter, we live in a world in which most of the interesting dependent variables are caused by more than one – often many more than one – independent variable. What problems does this pose for social science? It means that, when trying to establish whether a particular *X* causes a particular *Y*, we need to "control for" the effects of other causes of *Y* (and we call those other effects *Z*). If we fail to control for the effects of *Z*, we are quite likely to misunderstand the relationship between *X* and *Y* and make the wrong inference about whether *X* causes *Y*. This is the most serious mistake a social scientist can make. If we find that *X* and *Y* are correlated, but that, when we control for the effects of *Z* on both *X* and *Y*, the association between *X* and *Y* disappears, then the relationship between *X* and *Y* is said to be spurious. To return to our example about ice cream consumption (*X*) and homicide rates (*Y*), one obvious *Z* variable that might confound the relationship is "average monthly temperature." When it's warmer outside, people eat more ice cream. And when it's warmer outside, homicide rates rise. The association of both *X* and *Y* with *Z* can lead to the false appearance of a relationship between *X* and *Y*.

What does it mean to attempt to, as we have said, "control for" the effects of other variables? We'll eventually answer that question in two ways. Occasionally the control happens in the very design of the research plan; that possibility will be described in Chapter 4. More frequently, though, we will resort to statistical controls for these potentially confounding variables; that possibility, which happens far more frequently, will have to wait until Chapter 10.

### 3.2.1 Putting It All Together – Adding Up the Answers to Our Four Questions

As we have just seen, the process for evaluating a theoretical claim that *X* causes *Y* is complicated. Taken one at a time, each of the four questions in the introduction to this section can be difficult to answer decisively. But the challenge of evaluating a claim that *X* causes *Y* involves summing the answers to all four of these questions to determine our overall confidence about whether *X* causes *Y*. To understand this, think about the analogy that we have been using by calling these questions "hurdles." In track events that feature hurdles, runners must do their best to try to clear each hurdle as they make their way toward the finish line. Occasionally even the most experienced hurdler will knock over a hurdle. Although this slows them down and diminishes their chances of winning the race, all is not lost. If we think about putting a theory through the four hurdles posed by the preceding questions, there is no doubt our confidence will be greatest when we are able to answer all four questions the right way ("yes," "yes," "yes," "yes") and without reservation. As we described in the introduction to this section, failure to clear the first hurdle should give us pause. This is also the case if we find our relationship to be spurious. For the second and third hurdles, however, failure to clear them completely does not mean that we should discard the causal claim in question. Figure 3.1 provides a summary of this process. In the sections that follow, we will go through the process described in Figure 3.1 with a series of examples.

As we go through this process of answering the four questions, we will keep a **causal hurdles scorecard** as a shorthand for summarizing the answers to these four questions in square brackets. For now, we will limit our answers to "*y*" for "yes," "*n*" for "no," and "?" for "maybe." If a

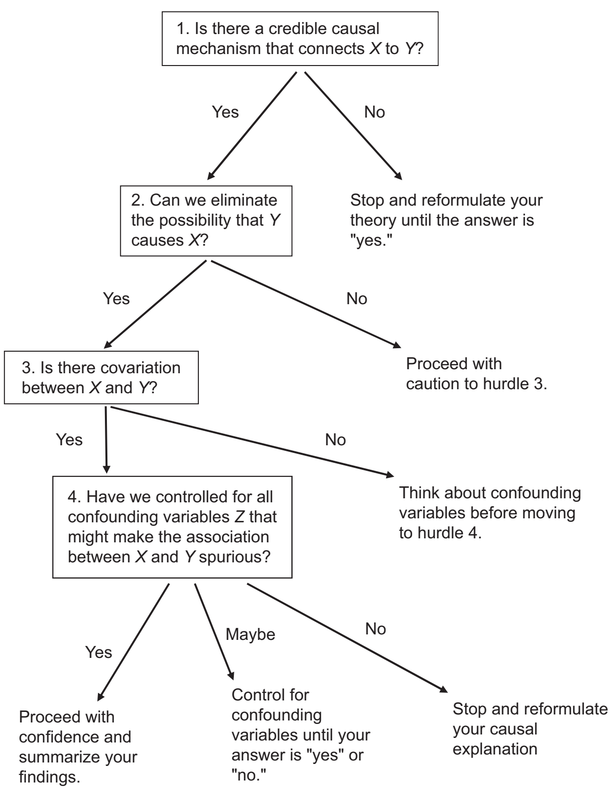


**Figure 3.1** The path to evaluating a causal relationship

theory has cleared all four hurdles, the scorecard would read [*y y y y*] and the causal claim behind it would be strongly supported. As we described above, these hurdles are not all the same in terms of their impact on our assessments of causality. So, for instance, a causal claim for which the scorecard reads [*n y y y*] could be thrown out instantly. But, a claim for which it reads [*y n y y*] would have a reasonable level of evidence in its favor.

### 3.2.2 Identifying Causal Claims Is an Essential Thinking Skill

We want to emphasize that the logic just presented does not apply merely to political science research examples. Whenever you see a story in the news, or hear a speech by a candidate for public office, or, yes, read a research article in a political science class, it is almost always the case that some form of causal claim is embedded in the story, speech, or article. Sometimes those causal claims are explicit – indented and italicized so that you just can't miss them. Quite often, though, they are harder to spot, and most of the time not because the speaker or writer is trying to confuse you. What we want to emphasize is that spotting and identifying causal claims is a thinking skill. It does not come naturally to most people, but it can be practiced.

In our daily lives, we are often presented with causal claims by people trying to persuade us to adopt their point of view. Advocacy and attempts at persuasion, of course, are healthy features of a vibrant democracy. The health of public debate, though, will be further enhanced when citizens actively scrutinize the claims with which they are presented. Take, for example, debates in the media about the merits of private school choice programs, which have been implemented in several school districts. Among the arguments in favor of such programs is that the programs will improve student performance on standardized tests. Media reports about the successes and failures of programs like this are quite common. For example, an article by Jay Mathews in the *Washington Post* discusses a study that makes the argument that:

> African American students in the District [of Columbia] and two other cities have moved ahead of their public school classmates since they transferred to private schools with the help of vouchers, according to a new study. . . . The study showed that those moving to private schools scored 6 percentile points higher than those who stayed in public schools in New York City, Dayton, Ohio, and the District. The effect was biggest in the District, where students with vouchers moved 9 percentile points ahead of public school peers.[7]

[7] Mathews, Jay. "Scores Improve for D.C. Pupils With Vouchers" *Washington Post*, August 28, (2000). A1.

Notice the causal claim here, which is: Participation (or not) in the school choice program ($X$) causes a child's test scores ($Y$) to vary. Often, the reader is presented with a bar chart of some sort in support of the argument. The reader is encouraged to think, sometimes subtly, that the differing heights of the bars, representing different average test scores for school choice children and public school children, means that the program *caused* the school choice children to earn higher scores. When we take such information in, we might take that nugget of evidence and be tempted to jump to the conclusion that a causal relationship exists. The key lesson here is that this is a premature conclusion.

Let's be clear: School choice programs may indeed cause students to do better on standardized tests. Our objective here is not to wade into that debate, but rather to sensitize you to the thinking skills required to evaluate the causal claim made in public by advocates such as those who support or oppose school choice programs. Evidence that students in school choice programs score higher on tests than do public school students is *one piece* of the causal puzzle – namely, it satisfies crossing hurdle 3 above, that there is covariation between $X$ and $Y$. At this point in our evaluation, our scorecard reads [? ? *y* ?]. And thus, before we conclude that school choice does (or does not) cause student performance, we need to subject that claim to all four of the causal hurdles, not just the third one.

So let's apply all four causal hurdles to the question at hand. First, is there a mechanism that we can use to explain how and why attending a particular type of school – public or a voucher-sponsored private school – might affect a student's test scores? Certainly. Many private schools that participate in voucher programs have smaller class sizes (among other benefits), and smaller class sizes can translate to more learning and higher test scores. *The answer to the first question is "yes"* [*y* ? *y* ?].

Second, is it possible that the causal arrow might be reversed – that is, can we rule out the possibility that test scores cause a person to participate or not participate in a school choice program? Since the test scores occur months or even years after the person chooses a school to attend, this is not possible. *The answer to the second question is "yes"* [*y y y* ?].

Third, is there a correlation between participation in the program and test scores? The article quoted above just noted that, in the three cities considered, there is – voucher school students scored higher on standardized tests than their public school peers. *The answer to the third question is "yes"* [*y y y* ?].

Finally, have we controlled for all confounding variables that might make the association between participation in the program and test scores spurious? Remember, a potentially confounding variable is simply a variable that is related to the independent variable and is also a cause of the dependent variable. So, can we think of something that is both related to the type of school a child attends and is also a likely cause of that child's test scores? Sure. The variable "parental involvement" is a natural candidate to be a $Z$ variable in this instance. Some children have highly involved parents – parents who read to their children, help them with homework, and take an active role in their education – while other children have parents who are much less involved. Highly involved parents are more likely than their uninvolved counterparts to learn about the existence of school choice programs in their cities, and are more likely to apply for such programs. (So $Z$ is almost surely related to $X$.) And highly involved parents are more likely to create high expectations among their children, and to instill in their children a sense that achievement in school is important, all of which probably translate into having children who score better on standardized tests. (So $Z$ is likely to be a cause of $Y$.) The key question then becomes: Did the study in question manage to *control for* those effects? We're a little ahead of the game here, because we haven't yet talked about the strategies that researchers employ to control for the effects of potentially confounding variables. (That task comes in Chapter 4.) But we hope you can see why controlling for the effects of parental involvement is so key in this particular situation (and in general): If our comparison of school choice children and public school children basically amounts to a comparison between the children of highly motivated parents and the children of poorly motivated parents, then it becomes very problematic to conclude that the difference between the groups' test scores was *caused by* the program. Without a control for parental involvement ($Z$), in other words, the relationship between school type ($X$) and test scores ($Y$) might be spurious. So, until we see evidence that this important $Z$ has been controlled for, our scorecard for this causal claim is [*y y y n*] and we should be highly suspicious of the study's findings. More informally, without such a control, the comparison between those sets of test scores is an unfair one, because the groups would be so different in the first place. As it happens, the article from the *Washington Post* that we mentioned did include a control for parental involvement, because the students were chosen for the program by a random lottery. We'll wait until Chapter 4 to describe exactly why this makes such a big difference, but it does.

The same process can be applied to a wide variety of causal claims and questions that we encounter in our daily lives. Does drinking red wine cause a reduction in heart disease? Does psychotherapy help people with emotional and relational problems? Do increases in government spending spur or retard economic growth? In each of these and many other examples, we might be tempted to observe a correlation between two variables and conclude that the relationship is causal. It is important for us to resist that temptation, and subject each of these claims to the more rigorous criteria that we are suggesting here. If we think about such evidence on its own in terms of our causal hurdles scorecard, what we have is [? ? *y* ?]. This is a reasonable start to the evaluation of a causal claim, but a pretty poor place to stop and draw definitive conclusions. Thinking in terms of the hurdles depicted in the scorecard, whenever someone presents us with a causal claim but fails to address each of the hurdles, we will naturally ask further questions and, when we do that, we will be much smarter consumers of information in our everyday lives.

**YOUR TURN: Does eating chocolate promote a healthy heart?**

Go read the following article: http://www.nytimes.com/2009/09/15/health/15choc.html

Based solely on what appears in the article, complete a causal hurdles scorecard about the claim that eating chocolate ($X$) causes a person to have less heart disease ($Y$).

An important part of taking a scientific approach to the study of politics is that we turn the same skeptical logic loose on scholarly claims about causal relationships. Before we can evaluate a causal theory, we need to consider how well the available evidence answers each of the four questions about $X$, $Y$, and $Z$. Once we have answered each of these four questions, one at a time, we then think about the overall level of confidence that we have in the claim that $X$ causes $Y$.

### 3.2.3 What Are the Consequences of Failing to Control for Other Possible Causes?

When it comes to any causal claim, as we have just noted, the fourth causal hurdle often trips us up, and not just for evaluating political rhetoric or stories in the news media. This is true for scrutinizing scientific research as well. In fact, a substantial portion of disagreements between scholars boils down to this fourth causal hurdle. When one scholar is evaluating another's work, perhaps the most frequent objection is that the researcher "failed to control for" some potentially important cause of the dependent variable.

What happens when we fail to control for some plausible other cause of our dependent variable of interest? Quite simply, it means that we have failed to cross our fourth causal hurdle. *So long as a reasonable case can be made that some uncontrolled-for Z might be related to both X and Y, we cannot conclude with full confidence that X indeed causes Y*. Because the main goal of science is to establish whether causal connections between variables exist, then failing to control for other causes of $Y$ is a potentially serious problem.

One of the themes of this book is that statistical analysis should not be disconnected from issues of research design – such as controlling for as many causes of the dependent variable as possible. When we discuss multiple regression (in Chapters 10, 11, and 12), which is the most common statistical technique that political scientists use in their research, the entire point of those chapters is to learn how to control for other possible causes of the dependent variable. We will see that failures of research design, such as failing to control for all relevant causes of the dependent variable, have statistical implications, and the implications are always bad. Failures of research design produce problems for statistical analysis, but hold this thought. What is important to realize for now is that good research design will make statistical analysis more credible, whereas poor research design will make it harder for any statistical analysis to be conclusive about causal connections.

**YOUR TURN: Exploring media reports of other social science studies**

The media outlet NPR has a regular series in its broadcasts that they call "Hidden Brain" which explores some of the subconscious forces that shape human beliefs and behavior. They have an active Twitter feed (*@HiddenBrain*) and an extensive series of podcasts.

Go visit their web site: http://www.npr.org/podcasts/510308/hidden-brain

Pick a podcast on a topic that interests you, and listen to hear how the host describes whether or not the relationships uncovered are *causal* or *spurious*. It takes practice!

## 3.3 WHY IS STUDYING CAUSALITY SO IMPORTANT? THREE EXAMPLES FROM POLITICAL SCIENCE

Our emphasis on causal connections should be clear. We turn now to several active controversies within the discipline of political science, showing how debates about causality lie at the heart of precisely the kinds of controversies that got you (and most of us) interested in politics in the first place.

### 3.3.1 Life Satisfaction and Democratic Stability

One of the enduring controversies in political science is the relationship between *life satisfaction in the mass public* and *the stability of democratic institutions*. Life satisfaction, of course, can mean many different things, but for the current discussion let us consider it as varying along a continuum, from the public's being highly unsatisfied with day-to-day life to being highly satisfied. What, if anything, is the causal connection between the two concepts?

Political scientist Ronald Inglehart (1988) argues that life satisfaction ($X$) *causes* democratic system stability ($Y$). If we think through the first of the four questions for establishing causal relationships, we can see that there is a credible causal mechanism that connects $X$ to $Y$ – if people in a democratic nation are more satisfied with their lives, they will be less likely to want to overthrow their government. *The answer to our first question is "yes"* [*y* ? ? ?]. Moving on to our second question: Can we eliminate the possibility that democratic stability ($Y$) is what causes life satisfaction ($X$)? We cannot. It is very easy to conceive of a causal mechanism in which citizens living in stable democracies are likely to be more satisfied with their lives than citizens living in nations with a history of government instability and less-than-democratic governance. *The answer to our second question is "no"* [*y n* ? ?]. We now turn to the third question. Using an impressive amount of data from a wide variety of developed democracies, Inglehart and his colleagues have shown that there is, indeed, an association between average life satisfaction in the public and the length of uninterrupted democratic governance. That is, countries with higher average levels of life satisfaction have enjoyed longer uninterrupted periods of democratic stability. Conversely, countries with lower levels of life satisfaction have had shorter periods of democratic stability and more revolutionary upheaval. *The answer to our third question is "yes"* [*y n y* ?]. With respect to the fourth question, it is easy to imagine a myriad of other factors ($Z$) that lead to democratic stability, and whether Inglehart has done an adequate job of controlling for those other factors is the subject of considerable scholarly debate. *The answer to our fourth question is "maybe"* [*y n y* ?]. Inglehart's theory has satisfactorily answered questions 1 and 3, but it is the answers to questions 2 and 4 that have given skeptics substantial reasons to doubt his causal claim.

**YOUR TURN: Other causes of democratic stability**

Draw a diagram with the $X$, $Y$, and $Z$ variables we identified in the Inglehart study on democratic stability.

Can you think of any other $Z$ variables that are likely to be correlated with $X$ (life satisfaction in a country) and are also likely to be a cause of $Y$ (longevity of democracy)?

### 3.3.2 Race and Political Participation in the United States

Political participation – the extent to which individual citizens engage in voluntary political activity, such as voting, working for a campaign, or making a campaign contribution – represents one of the most frequently studied facets of mass political behavior, especially in the United States. And with good reason: Participation in democratic societies is viewed by some as one measure of the health of a democracy. After decades of studying the variation in Americans' rates of participation, several demographic characteristics consistently stood out as being correlated with participation, including an individual's racial classification. Anglos, surveys consistently showed, have participated in politics considerably more frequently than either Latinos or African Americans. A comprehensive survey, for example, shows that during a typical election cycle, Anglos engaged in 2.22 "participatory acts" – such as voting, working for a campaign, making a campaign contribution, attending a protest or demonstration, and similar such activities – whereas comparable rates for African Americans and Latino citizens were 1.90 and 1.41 activities (see Verba et al., 1993, figure 1).

Is the relationship between an individual's race ($X$) and the amount that the individual participates in politics ($Y$) a causal one? Before we accept the evidence above as conclusively demonstrating a *causal* relationship, we need to subject it to the four causal hurdles. Is there a reasonable mechanism that answers the "how" and "why" questions connecting race and political participation? There may be reason to think so. For long portions of US history, after all, some formal and many informal barriers existed prohibiting or discouraging the participation of non-Anglos. The notion that there might be residual effects of such barriers, even decades after they have been eradicated, is entirely reasonable. *The answer to our first question is "yes"* [*y* ? ? ?]. Can we eliminate the possibility that varying rates of participation cause an individual's racial classification? Obviously, yes. *The answer to our second question is "yes"* [*y y* ? ?]. Is there a correlation between an individual's race and their level of participation in the United States? The data above about the number of participatory acts among Anglos, African Americans, and Latinos clearly show that there is a relationship; Anglos participate the most. *The answer to our third question is "yes"* [*y y y* ?]. Finally, have we controlled for all possible confounding variables $Z$ that are related to both race ($X$) and participation ($Y$) that might make the relationship spurious? Verba and his colleagues suggest that there might be just such a confounding variable: socio-economic status. Less so today than in the past, socio-economic status ($Z$) is nevertheless still correlated with race ($X$). And unsurprisingly, socio-economic status ($Z$) is also a cause of political participation ($Y$); wealthy people donate more, volunteer more, and the like, than their less wealthy counterparts. Once controlling for socio-economic status, the aforementioned relationship between race and political participation entirely vanishes (see Verba et al., 1993, table 8). In short, the correlation that we observe between race and political participation is spurious, or illusory; it is not a function of race, but instead a function of the disparities in wealth between Anglos and other races. Once we control for those socio-economic differences, the connection between race and participation goes away. *The answer to our fourth question is "no."* In this case, the effort to answer the fourth question actually changed our answer to the third question, moving our scorecard from [*y y y* ?] to [*y y n n*]. This is one of the important ways in which our conclusions about relationships can change when we move from a bivariate analysis in which we measure the relationship between one independent variable, $X$, and our dependent variable, $Y$, to a multiple variable analysis in which we measure the relationship between $X$ and $Y$ controlling for a second independent variable, $Z$. It is also possible for a lot of other things to happen when we move to controlling for $Z$. For instance, it is also possible for our scorecard to change from [*y y n n*] to [*y y y y*].

**YOUR TURN: Other causes of participation**

Draw a diagram with the $X$, $Y$, and $Z$ variables we identified in the Verba et al. (1993) study on political participation.

Can you think of any other $Z$ variables that are likely to be correlated with $X$ (racial classification) and are also likely to be a cause of $Y$ (political participation)?

### 3.3.3 Evaluating Whether "Head Start" Is Effective

In the 1960s, as part of the war on poverty, President Lyndon Johnson initiated the program "Head Start" to give economically underprivileged children a preschool experience that – the program hoped – would increase the chances that these poor children would succeed once they reached kindergarten and beyond. The program is clearly well intended, but, of course, that alone does not make it effective. Simply put: Does Head Start work? In this case, "work" would mean that Head Start could increase the chances that participants in the program would have better educational outcomes than nonparticipants.

It would be tempting, in this case, to simply compare some standardized test scores of the children who participated in Head Start with those who did not. If Head Start participants scored higher, then – voila! – case closed; the program works. If not, then not. But, as before, we need to stay focused on all four causal hurdles. First, is there some credible causal mechanism that would answer the "how" and "why" questions that connect Head Start participation ($X$) to educational outcomes ($Y$)? Yes. The theory behind the program is that exposure to a preschool environment that anticipates the actual school setting helps prepare children for what they will encounter in kindergarten and beyond. Head Start, in this sense, might help reduce discipline problems, and prepare students for reading and counting, among other skills. *The answer to our first question is "yes"* [*y* ? ? ?]. Is it possible, secondly, that the causal arrow might be reversed – in other words, can we rule out the possibility that educational outcomes ($Y$) could cause participation in Head Start ($X$)? Because testing would take place years after participation in the program, yes. *The answer to our second question is "yes"* [*y y* ? ?]. Is there an association between participation in the program and learning outcomes? Study after study has shown that Head Start participants fare better when tested, and have fewer instances of repeating a grade, than those who have no preschool experience. For example, a widely cited study shows that Head Start children do better on a vocabulary test suitable for young children than do students who have no preschool experience (Currie and Thomas, 1995). *The answer to our third question is "yes"* [*y y y* ?]. But, as was the case with the school-voucher example discussed previously, a potentially confounding variable – parental involvement ($Z$) – lurks nearby. Highly involved parents ($Z$) are more likely to seek out, be aware of, and enroll their children ($X$) in programs like Head Start that might benefit their children. Parents who are less involved in their children's lives are less likely to avail themselves of the potential opportunities that Head Start creates. And, as before, highly involved parents ($Z$) are likely to have positive effects on their children's educational outcomes. The key question, then, becomes: Do parental effects ($Z$) make the relationship between Head Start and later educational outcomes spurious? The aforementioned study by Currie and Thomas uses both statistical controls as well as controls in the design of their research to account for parental factors, and they find that Head Start has lasting educational effects only for Anglo children, but not for African American children (see their table 4). Again, that phrase "statistical controls" may not be quite as transparent as it will be later on in this book. For now, suffice it to say that these researchers used all of the techniques available to them to show that Head Start does, indeed, have positive effects for some, but not all, children. *The answer to our fourth question is a highly qualified "yes"* [*y y y y*].

## 3.4 WRAPPING UP

Learning the thinking skills required to evaluate causal claims as conclusively as possible requires practice. They are intellectual habits that, like a good knife, will sharpen with use.