1 The Scientific Study of Society

OVERVIEW

As instructors, we realize that most social science students are interested in the substance of the field and not in its methodology. We begin with a discussion of the goals of this book and why a scientific approach to the study of society is more interesting and desirable than a "just the facts" approach. In this chapter we provide an overview of what it means to study society scientifically. We begin with an introduction to how we move from causal theories to scientific knowledge, and a key part of this process is thinking about the world in terms of *models* in which the concepts of interest become variables that are causally linked together by theories. We then introduce the goals and standards of social science research that will be our rules of the road to keep in mind throughout this book. The chapter concludes with a brief overview of the structure of the book.

Doubt is the beginning, not the end, of wisdom
—Chinese proverb

1.1 SOCIAL *SCIENCE*?

"You must like people." "You must enjoy helping others." These are responses that students often hear after announcing that they are taking courses in sociology. In fact, although for some sociologists such factors may have provided initial motivation for pursuing the field, this is not the focus of modern sociology. Instead, sociology is about the scientific study of social phenomena. Perhaps like you, a great many of today's sociologists were attracted to this discipline as undergraduates because of intense interest in a particular social issue. Although we are often drawn into sociology based on such passions, the most respected sociological

research today is conducted in a fashion that makes it impossible to tell the personal views of the writer.

Many people taking their first sociological research course are surprised to find out how much science and, in particular, how much math are involved. We would like to encourage the students who find themselves in this position to hang in there with us – even if your answer to this encouragement is "But I'm only taking this class because they require it to graduate, and I'll never use any of this stuff again." Even if you never run a regression model after you graduate, having made your way through these materials should help you in a number of important ways. We have written this book with the following three goals in mind:

- *To help you consume academic social science research in your other courses.* One of the signs that a field of research is becoming scientific is the development of a common technical language. We aim to make the common technical language of social science accessible to you.
- *To help you become a better consumer of information.* In sociology and many other areas of scientific and popular communication, claims about causal relationships are frequently made. We want you to be better able to evaluate such claims critically.
- *To start you on the road to becoming a producer of scientific research on society.* This is obviously the most ambitious of our goals. In our teaching we often have found that once skeptical students get comfortable with the basic tools of social science, their skepticism turns into curiosity and enthusiasm.

To see the value of this approach, consider an alternative way of learning about society, one in which sociology courses would focus on "just the facts." Under this alternative, for example, a course on race and ethnicity might inform students that, in terms of residential segregation in urban areas in the U.S., African Americans are the minority group that is most highly segregated from whites, Asians are the least segregated, and Hispanics are intermediate between blacks and Asians. Moreover, data from the 2010 census indicate that, although the gradual decline in black–white segregation levels experienced in metropolitan areas between 1960 and 2000 continued into the first decade of the twenty-first century, segregation remains high. In 2010 the value of the black–white "index of dissimilarity" for all 367 metropolitan areas in the U.S. was 59 – meaning that, on average, 59 percent of either blacks or whites would have to change neighborhoods in order for the racial composition of each neighborhood to reflect the overall racial composition of the metropolitan area. The index values for Hispanic–white and Asian–white segregation were 48 and 41, respectively. Strikingly, in 2010 the average African

American urban dweller lived in a neighborhood that was 45 percent black and 35 percent white; the average white metropolitan inhabitant, on the other hand, lived in a neighborhood that was 75 percent white and only 8 percent black. Similarly, the neighborhood composition of the typical Hispanic metropolitan resident in 2010 was 46 percent Hispanic and 35 percent white, and that of the average Asian resident was 22 percent Asian and 49 percent white. In short, whites live in neighborhoods with only a minimal minority presence; blacks and Hispanics inhabit neighborhoods with high minority representation; and Asians live in neighborhoods where they are disproportionately represented, though with more whites than the neighborhoods of either African Americans or Hispanics (Logan and Stults, 2011).

As important as these facts are, however, the problem with a "just the facts" approach is that no attention is paid to the factors that cause segregation to be so high in the first place. The latter, more theoretical approach to the topic might focus on the primary explanations for persistent racial residential segregation in the U.S. discussed in the social science research literature: economic disparities among racial and ethnic groups that create differential access to quality neighborhoods; discrimination in the real estate, insurance, and mortgage-lending industries that make it difficult for minorities to rent or purchase homes in their neighborhoods of choice; and neighborhood residential preferences, not only of whites but of minority group members themselves (Massey and Denton, 1993). This approach helps us to better understand the dynamics of race and ethnicity in America today than a "just the facts" approach is capable of doing.

In this chapter we provide an overview of what it means to study society scientifically. We begin this discussion with an introduction to how we move from causal theories to scientific knowledge. A key part of this process is thinking about the world in terms of *models* in which the concepts of interest become **variables**[1] that are causally linked together by theories. We then introduce the goals and standards of social science research that will be our rules of the road to keep in mind throughout this book. We conclude this chapter with a brief overview of the structure of the book.

[1] When we introduce an important new term in this book, that term appears in boldface type. At the end of each chapter, we will provide short definitions of each bolded term that was introduced in that chapter. We discuss variables at great length later in this and other chapters. For now, a good working definition is that a variable is a definable quantity that can take on two or more values and is subject to change from one unit of analysis (e.g., a person, a neighborhood, or a city, to name a few common units of analysis) to another. An example of a variable is religion; researchers usually **measure** it in broad categories such as Protestant, Catholic, Jewish, Muslim, none, or other.

1.2 APPROACHING SOCIOLOGY SCIENTIFICALLY: THE SEARCH FOR CAUSAL EXPLANATIONS

I've said, I don't know whether it's addictive. I'm not a doctor. I'm not a scientist.

—Bob Dole, in a conversation with Katie Couric about tobacco during the 1996 U.S. presidential campaign

The question of "how do we know what we know" is, at its heart, a philosophical question. Scientists are lumped into different disciplines that develop standards for evaluating evidence. A core part of being a scientist and taking a scientific approach to studying the phenomena that interest you is always being willing to consider new evidence and, on the basis of that new evidence, change what you thought you *knew* to be true. This willingness to always consider new evidence is counterbalanced by a critical approach to the evaluation of new evidence that permeates the scientific approach. This is certainly true of the way that sociologists and other social scientists approach the study of society.

So what do social scientists do and what makes them scientists? A basic answer to this question is that, like other scientists, sociologists develop and test theories. A **theory** is a tentative conjecture about the causes of some phenomenon of interest. The development of **causal** theories about the social world requires thinking in new ways about familiar phenomena. As such, theory building is part art and part science. We discuss this in greater detail in Chapter 2, "The Art of Theory Building."

Once a theory has been developed, like all scientists, we turn to the business of testing our theory. The first step in testing a particular theory is to restate it as one or more testable hypotheses. A **hypothesis** is a theory-based statement about a relationship that we expect to observe framed in such a way that it can be empirically tested. For every hypothesis there is a corresponding **null hypothesis**. A null hypothesis is also a theory-based statement but it is about what we would observe if there were no relationship between an independent variable and the dependent variable. **Hypothesis testing** is a process in which scientists evaluate systematically collected evidence to make a judgment of whether the evidence favors their hypothesis or favors the corresponding null hypothesis. The process of setting up hypothesis tests involves both logical reasoning and creative design. In Chapter 3, "Evaluating Causal Relationships," we focus on the logical reasoning side of this process. In Chapter 4, "Research Design," we focus on the design part of this process. If a hypothesis survives rigorous testing, scientists start to gain confidence in that hypothesis rather than in the null hypothesis, and thus they also gain confidence in the theory from which they generated their hypothesis.

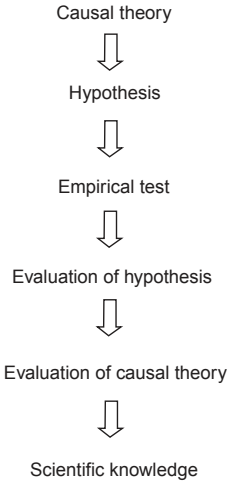

Figure 1.1 The road to scientific knowledge

Figure 1.1 presents a stylized schematic view of the path from theories to hypotheses to scientific knowledge.[2] At the top of the figure, we begin with a causal theory to explain our phenomenon of interest. We then derive one or more hypotheses about what our theory leads us to expect when we measure our concepts of interest (which we call variables – as subsequently discussed) in the real world. In the third step, we conduct **empirical** tests of our hypotheses.[3] From what we find, we evaluate our hypotheses relative to the corresponding null hypotheses. Next, from the results of our hypothesis tests, we evaluate our causal theory. In light of our evaluation of our theory, we then think about how, if at all, we should revise what we consider to be scientific knowledge concerning our phenomenon of interest.

A core part of the scientific process is skepticism. On hearing of a new theory, other scientists will challenge this theory and devise further tests. Although this process can occasionally become quite combative, it is a necessary component in the development of scientific knowledge. Indeed, a core component of scientific knowledge is that, as confident as we are in a particular theory, we remain open to the possibility that there is still a test out there that will provide evidence that makes us lose confidence in that theory.

It is important to underscore here the nature of the testing that scientists carry out. One way of explaining this is to say that scientists are *not* like lawyers in the way that they approach evidence. Lawyers work for a particular client, advocate a particular point of view (like "guilt" or "innocence"), and then accumulate evidence with a goal of proving their case to a judge or jury. This goal of *proving* a desired result determines their approach to evidence. When faced with evidence that conflicts with their case, lawyers attempt to ignore or discredit such evidence. When faced with evidence that supports their case, lawyers try to emphasize the applicability of the supportive evidence. In many ways, the scientific and legal approaches to evidence couldn't be further apart. Scientific confidence in a theory is achieved only after hypotheses derived from that theory have

[2] In practice, as we show in later chapters, the development of scientific knowledge is frequently much messier than this step-by-step diagram implies. Walter Wallace's classic, *The Logic of Science in Sociology* (1971), provides an excellent discussion of the complex interplay among these components of the scientific process.

[3] By "empirical" we simply mean "based on observations of the real world."

run a gauntlet of tough tests. At the beginning of a trial, lawyers develop a strategy to *prove* their case. In contrast, at the beginning of a research project, scientists will think long and hard about the most rigorous tests that they can conduct. A scientist's theory is never *proven* because scientists are always willing to consider new evidence that might contradict their theoretically based hypotheses.

The process of hypothesis testing reflects how hard scientists are on their own theories. As scientists evaluate systematically collected evidence to make a judgment of whether the evidence favors their hypothesis or favors the corresponding null hypothesis, they *always* favor the null hypothesis. Statistical techniques allow scientists to make probability-based statements about the empirical evidence that they have collected. You might think that, if the evidence was 50–50 between their hypothesis and the corresponding null hypothesis, the scientists would tend to give the nod to the hypothesis (from their theory) over the null hypothesis. In practice, though, this is not the case. Even when the hypothesis has an 80–20 edge over the null hypothesis, most scientists will still favor the null hypothesis. Why? Because scientists are very worried about the possibility of falsely rejecting the null hypothesis and therefore making claims that others ultimately will show to be wrong.

Once a theory has become established as a part of scientific knowledge in a field of study, researchers can build upon the foundation that this theory provides. Thomas Kuhn (1962) wrote about these processes in his famous book *The Structure of Scientific Revolutions*. According to Kuhn, scientific fields go through cycles of accumulating knowledge based on a set of shared assumptions and commonly accepted theories about the way that the world works. Together, these shared assumptions and accepted theories form what we call a **paradigm**. Once researchers in a scientific field have widely accepted a paradigm, they can pursue increasingly technical questions that make sense only because of the work that has come beforehand. This state of research under an accepted paradigm is referred to as **normal science**. When a major problem is found with the accepted theories and assumptions of a scientific field, that field will go through a revolutionary period during which new theories and assumptions replace the old paradigm to establish a new paradigm.

One of the more famous of these scientific revolutions occurred during the sixteenth century when the field of astronomy was forced to abandon its assumption that the Earth was the center of the known Universe. This was an assumption that had informed theories about planetary movement for thousands of years. In his book *De revolutionibus orbium coelestium* of 1543 (translated 2004 as *On the Revolutions of Heavenly Spheres*), Nicolaus Copernicus presented his theory that the Sun was the center of

the known Universe. Although this radical theory met many challenges, an increasing body of evidence convinced astronomers that Copernicus had it right. In the aftermath of this **paradigm shift**, researchers developed new assumptions and theories that established a new paradigm, and the affected fields of study entered into new periods of normal scientific research.

It may seem hard to imagine that the field of sociology has gone through anything that can compare with the experiences of astronomers in the sixteenth century. Indeed, Kuhn and other scholars who study the evolution of scientific fields of research have a lively and ongoing debate about where the social sciences, like sociology, are in terms of their development. The more skeptical participants in this debate argue that sociology is not sufficiently mature to have a paradigm, much less a paradigm shift. If we put aside this somewhat esoteric debate about paradigms and paradigm shifts, we can see an important example of the evolution of scientific knowledge about society from the study of public opinion in the United States.

In the 1940s the study of public opinion through mass surveys was in its infancy. Prior to that time, sociologists and political scientists assumed that U.S. voters were heavily influenced by presidential campaigns – and, in particular, by campaign advertising – as they made up their minds about the candidates. To better understand how these processes worked, a team of researchers from Columbia University set up an in-depth study of public opinion in Erie County, Ohio, during the 1944 presidential election. Their study involved interviewing the same individuals at multiple time periods across the course of the campaign. Much to the researchers' surprise, they found that voters were remarkably consistent from interview to interview in terms of their vote intentions. Instead of being influenced by particular events of the campaign, most of the voters surveyed had made up their minds about how they would cast their ballots long before the campaigning had even begun. The resulting book by Paul Lazarsfeld, Bernard Berelson, and Hazel Gaudet (1948), titled *The People's Choice*, changed the way that scholars thought about public opinion and political behavior in the United States. If political campaigns were not central to vote choice, scholars were forced to ask themselves what *was* critical to determining how people voted.

At first other scholars were skeptical of the findings of the 1944 Erie County study, but as the revised theories of politics of Lazarsfeld et al. were evaluated in other studies, the field of public opinion underwent a change that looks very much like what Thomas Kuhn calls a "paradigm shift." In the aftermath of this finding, new theories were developed to attempt to explain the origins of voters' long-lasting attachments to political parties in

the United States. An example of an influential study that was carried out under this shifted paradigm is Richard Niemi and Kent Jennings' seminal book from 1974, *The Political Character of Adolescence: The Influence of Families and Schools*. As the title indicates, Niemi and Jennings studied the attachments of schoolchildren to political parties. Under the pre-Erie County paradigm of public opinion, this study would not have made much sense. But once researchers had found that voters' partisan attachments were quite stable over time, studying them at the early ages at which they form became a reasonable scientific enterprise. You can see evidence of this paradigm at work in current studies of party identification and debates about its stability.

1.3 THINKING ABOUT THE WORLD IN TERMS OF VARIABLES AND CAUSAL EXPLANATIONS

So how do social scientists develop theories about society? A key element of this is that they order their thoughts about the social world in terms of concepts that scientists call *variables* and causal relationships between variables. This type of mental exercise is just a more rigorous way of expressing ideas about society that we hear on a daily basis. You should think of each variable in terms of its *label* and its *values*. The **variable label** is a description of what the variable is, and the **variable values** are the categories or metrics in which the variable occurs. So, if we're talking about the variable that reflects an individual's age, we could simply label this variable "Age" and the metric in which this variable occurs would typically be years, although months, days, or even hours might apply in some situations.

It is easier to understand the process of turning concepts into variables by using an example of an actual theory. For instance, if we're curious about the factors that shape people's attitudes toward immigrants, one possibility is that attitudes will be more positive when the U.S. economy is perceived as being relatively healthy. If we restate this in terms of a social science theory, the perceived state of the economy becomes the **independent variable** and attitudes toward immigrants becomes the **dependent variable**. One way of keeping the lingo of theories straight is to remember that the value of the "dependent" variable "depends" on the value of the "independent" variable. Recall that a theory is a tentative conjecture about the causes of some phenomenon of interest. In other words, a theory is a conjecture that the independent variable is causally related to the dependent variable; according to our theory, change in the value of the independent variable *causes* change in the value of the dependent variable.

This is a good opportunity to pause and try to come up with your own causal statement in terms of an independent and dependent variable; try filling in the following blanks with some variables that interest you:

_____ causes _____

Sometimes it's easier to phrase causal propositions more specifically in terms of the values of the variables that you have in mind. For instance,

higher _____ causes lower _____

or

higher _____ causes higher _____

Once you learn to think about the world in terms of variables, you will be able to produce an almost endless slew of causal theories. In Chapter 4 we will discuss at length how we design research to evaluate the causal claims in theories, but one way to initially evaluate a particular theory is to think about the **causal explanation** behind it. The causal explanation behind a theory is the answer to the question, "*Why* do you think that this independent variable is causally related to this dependent variable?" If the answer is reasonable, then the theory has possibilities. In addition, if the answer is original and thought provoking, then you may really be on to something. Let's return now to our working example in which the perceived state of the economy is the independent variable and attitudes toward immigrants is the dependent variable. The causal explanation underlying this theory is our belief that the perceived state of the economy is *causally related* to views of immigrants *because* economic insecurity creates fear of competition for jobs and other resources among members of the host society. As a result, when the economy is thought to be performing well, attitudes toward immigrants will be more positive. When the economy is thought to be performing poorly, attitudes will be more negative. If we put this in terms of the preceding fill-in-the-blank exercise, we could write:

perceptions of economic performance cause
attitudes toward immigrants.

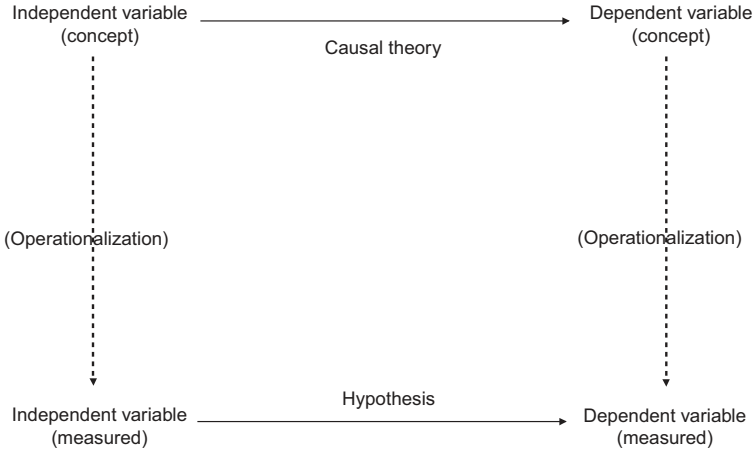

Figure 1.2 From theory to hypothesis

or, more specifically, we could write:

perceptions of higher economic performance cause more positive views of immigrants.

For now we'll refer to this theory, which has been widely advanced and tested by sociologists and political scientists alike, as "the theory of economic threat."

To test the theory of economic threat, we need to derive from it one or more testable hypotheses. Figure 1.2 provides a schematic diagram of the relationship between a theory and one of its hypotheses. At the top of this diagram are the components of the causal theory. As we move from the top part of this diagram (Causal theory) to the bottom part (Hypothesis), we are moving from a general statement about how we think the world works to a more specific statement about a relationship that we expect to find when we go out in the real world and measure (or **operationalize**) our variables.[4]

At the theory level at the top of Figure 1.2, our variables do not need to be explicitly defined. With the economic threat example, the independent variable, labeled "perceptions of economic performance," can be thought of as a concept that ranges from values of very strong to very poor. The dependent variable, labeled "attitudes toward immigrants," can be thought of as a concept that ranges from values of very positive to very negative. Our causal theory is that a stronger economic performance causes more

[4] Throughout this book we will use the terms "measure" and "operationalize" interchangeably, as is common practice in current social science literature.