4

Elements of Research Design

All social research focuses on a particular topic and addresses specific research questions. We begin this chapter by considering the social forces and personal motives that affect topic selection. We then consider three main concerns as a researcher narrows in on researchable questions: (1) what entities (e.g., individual people, groups, formal organizations, nations) are to be studied, (2) what aspects or characteristics of these entities are of interest, and (3) what kinds of relationships among the characteristics are anticipated. By considering these key elements of research design, we can grasp the language of social research and what it means to state the problem in researchable terms.

Origins of Research Topics

The starting point for research is the selection of a topic. Once a topic is chosen and the research question is set, we can consider guidelines for conducting research that will generate the most valid data and the most definitive answers. Given that anything that is "social" and "empirical" could be the subject of social research, the variety of potential topics is nearly endless. So, *how* are specific topics likely to emerge in the social sciences? We have identified five factors that explain the origin of most topics.

1. *The structure and state of the scientific discipline.* With the scientific goal of advancing knowledge, most researchers select topics suggested by the ongoing development of theory and research in their particular fields of study. The organization of disciplines casts the framework for topic selection. Social psychologists, for example, divide their discipline with respect to various forms of social behavior, such as aggression, altruism, interpersonal

attraction, and conformity, which act as organizing themes or areas of research interest. Similarly, sociologists frequently study aspects of various institutions, like religion, politics, education, and the family, around which the discipline is organized. As knowledge in an area develops, inconsistencies and gaps in information are revealed that researchers attempt to resolve.

2. *Social problems.* The focus and development of the social sciences are intimately related to interest in basic problems of the "human condition." Historically, this has been a major source of research topics, especially in sociology. The most eminent sociologists of the nineteenth and early twentieth centuries—people like Emile Durkheim, Karl Marx, Max Weber, and Robert Park—concerned themselves with problems emanating from great social upheavals of their day, such as the French and Industrial revolutions and massive foreign immigration to the United States. The problems wrought by these events—alienation, deviance, urban crowding, racism, and many others—have remained a major focus of the discipline. Indeed, many people are attracted to the social sciences because of their perceived relevance to social problems.

3. *Personal interests of the researcher.* Carrying out a research project, with its inevitable complications, obstacles, and demands for time and money, requires considerable interest and commitment on the part of the investigator. What sustains this interest more than anything else are highly personal motivations for doing research on a particular topic. Thus, an investigator may choose a topic not only because it is considered theoretically important, novel, or researchable but also because it stimulates his or her interest. According to the social psychologist Zick Rubin (1976:508–9), one reason for his embarking on the study of romantic love was that he was "by temperament and avocation, a songwriter. Songwriters traditionally put love into measures." And so he set out to find a way of measuring love scientifically. In a similar fashion, it is not surprising that members of particular groups often have pioneered research on those groups; for example, women have led the way in research on women and African Americans in research on blacks (King, Keohane, and Verba, 1994:14).

4. *Social premiums.* There are also powerful social determinants of topic selection. Through the availability of supporting funds, the prestige and popularity of the research area, and pressures within the discipline and within society, social premiums are placed on different topics at different times. Typically, these premiums will reinforce one another, with the social climate affecting funding, which in turn affects prestige. This was certainly true of the space program in the 1960s. Today, in the social sciences, the aging of the population as a whole has raised interest and support for research on the elderly, just as it has caused a dramatic increase in federal expenditures on and services available to older people. Similarly, in the 1970s the women's movement spurred a dramatic increase in research on gender issues that has continued into the twenty-first century.

5. *Practical considerations.* An overriding concern in any research project is cost. Research requires time, money, and personnel. Limitations on these resources, as well as other practical considerations such as the skill of the researcher and the availability of relevant data (see Chapter 12), will shape both the nature and the scope of the problem that the researcher can pursue.

The choice of any given research topic may be affected by any or all of the factors mentioned. Consider a study by Beckett Broh (2002), which examined the effects of participation in extracurricular activities on high school academic achievement. Using data from a national survey of high school students, Broh's study was designed to determine who benefits from participating in sports and other school activities and why. The study continued a line of inquiry on the impact of the extracurriculum of theoretical interest to sociologists of education and sport. Social scientists, school officials, and the general public have long debated whether sport, in particular, builds character and has positive educational benefits. Given the costs of extracurricular programming, especially school sports, and the public concern about boosting academic achievement, Broh's study also had important practical implications. Finally, the topic was of special personal interest to Broh. She had been a high school athlete, and after her collegiate athletic career was cut short by an injury, she turned to coaching middle and high school basketball in Michigan and Ohio. These experiences naturally sparked her interest in sport and education as a PhD student in sociology. And when her mentor suggested that she look at the questions in the National Educational Longitudinal Study, she found the means of testing some of her ideas about the impact of athletes' network of relationships on their academic achievement.

Once a general topic has been chosen, it must be stated in researchable terms. This involves translating the topic into one or more specific questions or hypotheses. We'll discuss the process of moving from general topics to specific questions later. First, it is important to understand that the formulation of a researchable problem or question boils down to deciding what *relationships* among what *variables* of what *units* are to be studied. We will now turn our attention to these important terms.

Units of Analysis

The entities (objects or events) under study are referred to in social research as **units of analysis**. Social scientists study a variety of units (sometimes called "elements" or "cases"). These include individual people; social roles, positions, and relationships; a wide range of social groupings such as families, organizations, and cities; and various social artifacts such as books, periodicals, documents, and even buildings. Ordinarily, the unit of analysis is easily identified. The unit is simply what or who is to be described or analyzed. For example, a researcher wanting to determine whether larger organizations (in terms of the number of employees) have more bureaucratic rules and regulations than smaller ones would treat the *organization* as the relevant unit and would gather information on the size and bureaucratic complexity of different organizations. In Broh's study of the impact of extracurricular activities on academic achievement, the unit of analysis was *individuals* or, more precisely, individual high school students.

In much social research the individual person is the unit of analysis. Social scientists are less interested in individual differences, however, than in the impact of the social context on people in general, the social relationships that individuals form, and large-scale social processes. To analyze "the social" often requires units beyond individual human beings. Consider Fred Markowitz's (2006) study of psychiatric deinstitutionalization. Before the

1960s substantial numbers of mentally ill persons were treated in large, publicly funded hospitals. But as a result of developments in psychiatric drugs, stricter standards for involuntary commitment, and federal funding cuts, the number of resident patients in state and county psychiatric hospitals sharply declined over the next few decades. Studies of *individuals* seem to indicate that this deinstitutionalization movement increased the number of mentally ill who are arrested or imprisoned. For Markowitz, this raised a question about social control: If the behavior of mentally ill persons, who are more likely than others to be both perpetrators and victims of crime, cannot be managed by the mental health–care system, will the mentally ill tend to end up in the criminal justice system? Mental health-care and criminal justice *systems* refer to *social* units; social systems are aspects of communities or nations, not individuals. Markowitz thus chose to analyze cities. Examining the number of psychiatric beds in hospitals and the crime and arrest rates in 81 U.S. cities, he found that the lower a city's capacity to place the mentally ill in public psychiatric hospitals, the higher the city's crime and arrest rates.

Data on individuals also may be inadequate for investigating social and cultural change, especially in the distant past. At mid-twentieth century, the sociologist David Riesman (1950) published an influential book, *The Lonely Crowd*, in which he theorized a general trend toward "other-directedness." Because of changes wrought by the Industrial Revolution, such as the expansion of white-collar and service jobs, increasing material abundance, and the development of mass media of communication, people's actions were becoming less motivated by intrinsic values and more influenced by the actions of others. But how can one study long-range trends in individual motivation when it is impossible to analyze individuals from the past? One social scientific solution is to rely on various social artifacts such as the media and to assume that their visual or verbal imagery reflects the individual values and behavior of direct interest. To test Riesman's theory, for example, researchers have chosen advertisements in mass-circulation magazines as their units of analysis. One study analyzed more than 4000 ads in three women's magazines (*McCall's*, *Ladies' Home Journal*, and *Good Housekeeping*) for the period 1890–1984 (Zinkhan and Hayes, 1989). Results showed that the advertising appealed increasingly over time to the standards of others (other-directedness).

Aggregate Data

One way of identifying the unit of analysis in reported research is to determine the properties or traits that are being described. Individual people have traits such as age, race, and gender; have attitudes and opinions; and act in certain ways, such as saying that they voted in the last civic election, committed a felony, or did volunteer work last week. Only social aggregates such as groups, organizations, cities, and nations have a size, average age, percentage who voted, or crime rate. It can be confusing, however, when individual-level data are combined to describe a social unit to which the individuals belong.

Information about one set of units that is statistically combined to describe a larger social unit is called **aggregate data**. Each year, *U.S. News and World Report* uses aggregate data to rank America's best colleges: They gather information about individual students at each college to characterize the college as a whole. Among their indicators of academic quality, for example, is each school's graduation rate, which is calculated by dividing the number of enrollees in a given class into the number who graduate within 6 years of their enrollment. Other aggregate measures include average SAT or ACT test scores; the proportion of enrolled freshmen who graduated in the top 10 percent of their high school classes; and the acceptance rate, the ratio of the number of students admitted to the number who apply.

In *U.S. News* rankings, the unit of analysis is colleges; that is what is being described and compared. But the use of aggregate data does not always imply a collective unit of analysis. In another study of adolescents' extracurricular participation, Andrew Guest and Barbara Schneider (2003) wondered whether the effect of such activities depended on the social context of the school and community. For example, sports may have a greater effect on academic achievement in lower-class communities, which value sports as paths to financial gain, than in upper-class communities, which value sports for their health and aesthetic benefits. Guest and Schneider's unit of analysis, as in Broh's study, was individual high school students; however, they created aggregate measures to study the impact of the community and school context. Their measure of the community's socioeconomic class was based largely on the average income and education of the neighborhoods where the students lived, and a school context measure consisted of the percentage of students from each student's high school who went on to 4-year colleges after graduation.

Thus, when information about individuals is aggregated to describe groups or collectivities, the unit of analysis may be either the individual or the group. How can you tell? In their analyses, researchers ordinarily compare a number of instances of a particular unit—for example, a number of individuals, a number of cities, or a number of colleges. To identify the unit of analysis, therefore, ask yourself what is being described in each instance and what sorts of units are being compared—individuals or groups. If the aggregate data are used to compare different groups, the unit of analysis is the group. If the data are used to compare individuals who belong to different groups, the unit is the individual.

One reason it is important to identify accurately the unit of analysis is that confusion over units may result in false conclusions about research findings. Generally speaking, assertions should be made only about the particular unit under study. (Actually, assertions should be even more circumscribed than this, as we will see in Chapter 6.) To draw conclusions about one unit on the basis of information about another is to risk committing a logical fallacy.

Ecological Fallacy

The most common fallacy involving the mismatching of units is the **ecological fallacy** (Robinson, 1950). This occurs when relationships between properties of groups or geographic areas are used to make inferences about the individual behaviors of the people within those groups or areas. This is similar to what logicians call the "fallacy of division": assuming that what holds true of a group also is true of individuals within the group. Knowing that Sally attended a college whose students had relatively low average SAT scores, you would commit this fallacy if you assumed that Sally herself had low SAT scores.

Political analysts who use aggregate data from elections to study individual voting behavior are particularly susceptible to the ecological fallacy. Suppose, for example, that a researcher wanted to know whether registered Democrats or Republicans were more likely to support an Independent Party candidate in a city election but that the only information available was the percentage of votes the candidate received and the percentages of Republican and Democratic voters *in each precinct*. In short, the researcher wants to draw conclusions about *individual voters* but has collective information only about *precincts*. Knowing that the candidate received a relatively larger number of votes in precincts with greater percentages of Republicans does not permit the conclusion that Republican voters were more likely to support the candidate than were Democratic voters. In fact, it is plausible that Democrats in predominantly Republican precincts were more likely to support the Independent Party candidate than were Democrats in other precincts. Because the unit of analysis here is the precinct and not the individual voter, we simply do not know how individual voters within each precinct voted.

At one time, social scientists frequently performed ecological analyses such as the one above. For example, criminologists analyzed crime and delinquency rates in relation to other characteristics of census tracts to draw conclusions about characteristics of individual criminals and delinquents. A typical erroneous conclusion might be that foreign-born persons commit more crimes than native-born persons because the crime rate is higher in areas with greater proportions of foreigners. But such a conclusion is clearly unwarranted because we do not know who actually committed the crimes—foreign or native-born persons. Similarly, Durkheim's classic study of suicide was subject to the ecological fallacy by inferring that Protestants commit more suicides than Catholics from the observation that suicide rates were higher in predominantly Protestant nations than in predominantly Catholic ones.[1]

It is not always wrong to draw conclusions about individual-level processes from aggregate or group-level data. Social scientists have identified conditions under which it is reasonable to make such inferences (e.g., Firebaugh, 1978), but it is often difficult to determine whether these conditions are met. The implications of the ecological fallacy are clear: Carefully determine the units about which you wish to draw conclusions and then make sure that your data pertain to those units. If you are interested in individuals but only aggregate data are available, then draw conclusions tentatively, recognizing the possibility of an ecological fallacy.

KEY POINT

If data describe social units (e.g., schools), one must be cautious in drawing conclusions about individuals (students within the schools).

Variables

Although researchers observe units of analysis, it is relationships among characteristics of units that are of primary interest. Characteristics of units that vary, taking on different values, categories, or attributes for different observations, are called **variables**. Variables may vary over cases, over time, or over both cases and time. For example, among individuals, any set of characteristics that may differ for different people, such as age (range of years), gender (male and female), and marital status (single, married, divorced, widowed, etc.), is a variable. And for an individual, any characteristic that may vary from one time period to the next, such as age, level of education (first grade, second grade, etc.), and income (dollars earned per year), is a variable.

It is not unusual to see some confusion between variables and the attributes or categories of which they consist. "Gender" is a variable consisting of the categories male and female; "male" and "female" by themselves are not variables but simply categories that distinguish persons of different gender. Likewise, "divorced" and "Republican" are not variables but categories of the variables "marital status" and "political party affiliation," respectively. To keep this distinction clear, note that any term you would use to describe yourself or someone else (e.g., sophomore, sociology major) is an attribute or category of a variable (academic class, major).

KEY POINT
Social scientists study *relationships* among *variables*; variables depict and differentiate *units of analysis.*

You are now in a position to test your understanding of the concepts of unit of analysis and variable. Table 4.1 is designed to help you do this. The first column presents eight research questions or hypotheses; the second and third columns identify the relevant units of analysis and variables. Read the first column while covering up the second and third columns and record the unit of analysis and variables on a separate sheet of paper. Then check your answers by examining the full table.

Types of Variables

Social scientists find it necessary to classify variables in several ways. One type of classification is necessitated by the complexity of social phenomena. For any given research problem, a researcher can observe and measure only a few of the many potentially relevant properties. Those variables that are the object of study—part of some specified relationship—are called **explanatory variables**, and all other variables are **extraneous** (Kish, 1959).

There are two types of explanatory variables: dependent and independent. The **dependent variable** is the one the researcher is interested in explaining and predicting. Variation in the dependent variable is thought to depend on or to be influenced by certain other variables. The explanatory variables that do the influencing and explaining are called **independent**. If we think in terms of cause and effect, the independent variable is the presumed cause and the dependent variable the presumed effect. Independent variables are also called "predictor variables" because their values or categories may be used to predict the values or categories of dependent variables. For example, when Broh (2002) studied the impact of extracurricular involvement on academic achievement, her independent variable consisted of whether students participated in specific school activities such as interscholastic sports and her dependent variable was level of academic achievement. One research question was whether sports participation explained (or predicted) differences in academic achievement.

TABLE 4.1. Sample Research Questions, Units of Analysis, and Variables

Research question/hypothesis	*Unit of analysis*	*Variables*
[WHAT ONE WANTS TO KNOW]	[WHAT ENTITIES ARE DESCRIBED AND COMPARED]	[WITH RESPECT TO WHAT CHARACTERISTICS]
Are older people more afraid of crime than younger people?	Individuals	Age, fear of crime
The greater the growth of air passenger traffic at a city's airport, the greater the economic growth.	Cities	Growth of air traffic, economic growth
The higher the proportion of female employees, the lower the wages in nineteenth-century factories.	Factories	Proportion of employees who are female, average wage
Does economic development lower the birth rate?	Nations	Level of economic development, birth rate
The longer the engagement period, the longer the marriage.	Couples (dyads)	Length of engagement, marriage duration
Fan support in the NBA, as measured by attendance, is not related to the proportion of black players on the team.	NBA teams	Racial composition of the team, average attendance
Comic strips introduced in the 1930s were more likely to emphasize powerful heroes than strips introduced in the 1920s.	Comic strips	Whether main characters in strip were powerful, when comic strips were introduced (1920s or 1930s)
When television sports commentators make reference to individual competitors, they are more likely to use first names for female athletes than for male athletes.	References to competitors	Gender of athlete, whether athlete is identified by first name

Research studies in the social sciences often involve several independent variables and sometimes more than one dependent variable. Also, a variable is not intrinsically independent or dependent. An independent variable in one study may well be a dependent variable in another, depending on what the researcher is trying to explain. Finally, it is conventional in mathematics and science for the letter X to symbolize the independent variable and for the letter Y to represent the dependent variable. This is a practice we shall follow in the remainder of the book.

Extraneous variables, which are not part of the explanatory set, may be classified in two important ways. First, in relation to specific independent and dependent variables, they may be **antecedent** or **intervening**. An antecedent variable occurs prior in time to both the independent and the dependent variable; a variable is intervening if it is an effect of the independent variable and a cause of the dependent variable. Antecedent variables in Broh's study were parents' income and a student's race and gender; each of these variables may affect both extracurricular involvement and academic achievement. An intervening variable was students' self-esteem. Extracurricular involvement could affect self-esteem, which in turn could affect a student's academic performance. Figure 4.1 depicts these examples of antecedent and intervening variables. Each arrow in the figure represents causal direction. Thus, "Parents' income → Extracurricular involvement" means that parents' income influences or causes extracurricular involvement, and the absence of an arrow means that one variable does not cause another.

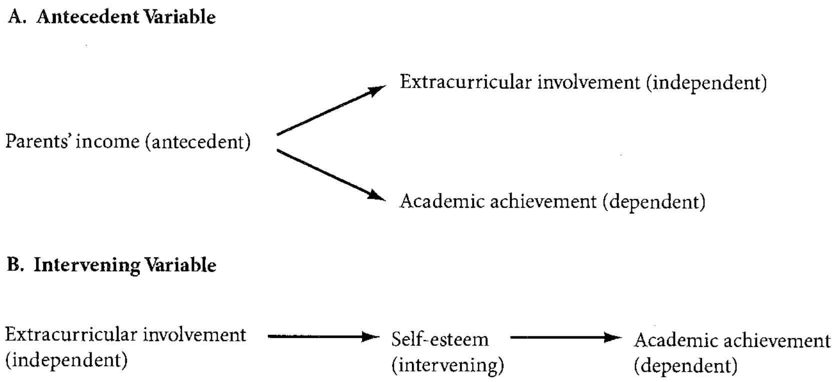


FIGURE 4.1. **Antecedent and intervening variables.**

Extraneous variables also may be categorized as controlled or uncontrolled. Controlled or, more commonly, **control variables** are held constant, or prevented from varying, during the course of observation or analysis. This may be done to limit the focus of the research or to test hypotheses pertaining to specific subgroups—for example, all males or all males younger than 18 years of age. Basically, the value or category of a control variable remains the same for a given set of observations. Several techniques for holding variables constant are discussed at length in the following chapters. Examples include selecting only individuals of the same age and gender, observing groups of the same size, creating uniform laboratory conditions or social settings in which to observe people or groups, and statistically controlling for specific attributes.

Whenever a variable is held constant in research, that variable cannot account for (or explain) any of the variation in the explanatory variables. Suppose, for example, that you wanted to explain differences (i.e., variation) between people in their level of aggression. If you controlled for gender by studying only males, then the variable "gender" could not account for any of the observed variation in aggression. Holding variables constant thus simplifies complex social situations. It is a means of ruling out variables that are not of immediate interest but that might otherwise explain part of the phenomenon that the investigator wishes to understand. One aim of efficient research design is to identify potentially relevant extraneous variables to measure and control as many as is feasible. Broh measured numerous extraneous variables, including parents' income and students' race, gender, and self-esteem. This enabled her to statistically control and test the effects of these variables on the relationship between extracurricular activities and academic achievement.

Another important distinction is made between quantitative and qualitative variables. This is a crude way of pointing to some significant differences in the numerical representation of variable categories. We make more precise distinctions along these lines in the next chapter, but for now we appeal to the common understanding of quantitative and qualitative. A variable is **quantitative** if its values or categories consist of numbers and if differences between categories have numerical value. For example, the variable "income," measured in dollars, signifies a quantitative difference—a certain number of dollars—between people with different incomes. **Qualitative variables** have discrete categories, usually designated by words or labels, and nonnumerical differences between categories. For example, the variable gender, consisting of the discrete categories male and female, makes a categorical, but not a numerical, distinction between people.[2]

Relationships

Social scientists' ultimate objective is to make sense of the social world by discovering enduring relationships among phenomena. Much research is therefore aimed directly at developing and testing relationships. There are other, more immediate purposes for which research is conducted, such as exploration and description, which we discuss later in the chapter. Yet, even when a researcher's goal is discovery or description, research findings will depend to a large extent on what particular *relationships* are anticipated. Research is not like the kind of fishing trip in which you drop your line anywhere, hoping you will catch something. Investigators do not make random observations. Whatever their goals, they must decide what to observe or ignore, how to go about making their observations, and how to interpret them. Such decisions inevitably are based on a researcher's expectations about how variables are related to one another.

Consider an investigator who wants to study "everything" about families but "let the facts speak for themselves" (Batten, 1971:9–11). Without *any* expectations, how is the researcher to decide what constitutes a "family"? Should first cousins be included even if they live far away or seldom keep in touch? Should a family include all persons living in the same household, no matter how remote the blood tie (Batten, 1971)? When collecting data, should the researcher include the hat size and shoe size of each family member? Such considerations ultimately depend on the social researcher's *expectations* about salient properties and *relationships* regarding families.

Facts never speak for themselves. Not only which facts are sought but also how they are interpreted depends again on anticipated relationships—on what particular answers the researcher expects the data or "facts" to provide. A classic everyday example (although likely apocryphal) of how one can draw very different conclusions from the same facts is "contrasting newspaper coverage of a two-car race between Soviets and Americans during the Cold War. An American newspaper described the race this way: 'American car beats out Soviet competitor.'" By contrast, a Russian newspaper described the exact same facts "differently: 'Soviet car finishes second; American car is next-to-last'" (D. Murray, Schwartz, and Lichter, 2001:86). Similarly, pundits and politicians often refer to the number of jobs "created" in the United States during a president's term of office. The most frequently cited numbers are estimates of employment drawn from a Bureau of Labor Statistics survey of businesses and government agencies. Conclusions drawn about these data, however, must take into account factors such as population growth and women's level of participation in the labor market, and they depend on assumptions about who is responsible for job growth (e.g., Congress, the president, global economy). Without considering the complex set of relationships among factors affecting employment, different conclusions may be drawn.

Given that all research carries expectations about the nature of what is being investigated, anticipated relationships and guiding theoretical explanations should always be identified as far as possible. Beckett Broh's (2002) selection of variables for her study was guided largely by three theoretical explanations of the link between sports participation and educational achievement: (1) developmental theory—sports teaches skills, such as a strong work ethic and perseverance that help students achieve in the classroom; (2) the leading crowd hypothesis—athletes gain social status and membership in a peer group of disproportionately college-oriented high achievers, which facilitates higher academic performance; and (3) social capital theory—sports create strong social ties among students, parents, and teachers that provide educational benefits. Broh derived the social capital model in part from her experience as a basketball coach; she noted that her players tended to spend more time with their parents and had better relationships with them than other students.

Thus far we have relied on the reader's intuitive grasp of the term "relationship." Everyone has experienced relationships at some point. One might think of relationships among kin; "serious" relationships, as between lovers; and relationships among the members of teams and work groups. We also have a sense of a relationship when one event regularly precedes or follows the occurrence of another, for example, when the appearance of dark clouds is regularly followed by rain. All such relationships have two features. First, they always involve two or more entities: persons, objects, or events—such as parent and child, leader and follower, or clouds and rain. Second, the pairs or combinations of things usually occur together and change together; thus, the appearance of one thing signals the appearance of the other and the absence of one implies the absence of the other. For example, by definition, every parent has a child and every child has (or had) a parent; we cannot have one without the other. Also, by observation we know that certain kinds of clouds produce rain and that without clouds it cannot rain.

KEY POINT

Facts or data do not speak for themselves; they must be interpreted based on anticipated and predicted relationships.

The kind of relationships with which social scientists are concerned, **relationships among variables**, have these same two features. Two or more variables are related, or form a relationship, to the extent that changes in one are accompanied by systematic changes in the other(s). Since the manner in which the variables change or vary together will depend on whether the variables are qualitative or quantitative, we will consider the nature of relationships separately for each of these types of variables.

Relationships among Qualitative Variables

Consider the two qualitative variables race and political party affiliation. If two individuals have the same party affiliation, say Democrat, then the category of this variable does not change as we look from one individual to the other. If they have different affiliations, say one is a Democrat and the other is a Republican, then the category of the variable does change. Assuming similar statements about race, we would say that a relationship exists between race and political party affiliation if a comparison of a pair of individuals reveals that a change in race is accompanied by a change in party affiliation. More generally, the basic

Table 4.2. Varying Degrees of Association between Two Qualitative Variables: Race and Political Affiliation

A. Perfect association	*White*	*Black*	*Total*
Democrat	0	10	10
Republican	10	0	10
Total	10	10	20
B. Moderate association	*White*	*Black*	*Total*
Democrat	3	7	10
Republican	7	3	10
Total	10	10	20
C. No association	*White*	*Black*	*Total*
Democrat	5	5	10
Republican	5	5	10
Total	10	10	20

idea of a relationship, or association, between qualitative variables can be incorporated into two assertions (Leik, 1972:26): (1) if one variable changes, the other variable changes, and (2) if one variable does not change, the other does not change. If these two assertions were true for all pairs of cases, the result would be a perfect one-to-one correspondence between the categories of one variable and the categories of the other. Table 4.2A depicts a perfect relationship between race and political party affiliation among 20 individuals.

In actual research, of course, we never see such perfect associations. To a researcher, therefore, the important question is not whether a given pair of variables is perfectly associated but how strongly they are related. That is, how closely do the data approximate a perfect association between variables? Statistical techniques to assess this are called "measures of association." These techniques, as well as the concept of "strength of relationship," are best understood if assertions about concomitant changes in variables are treated as predictions. Table 4.2A suggests two predictions: (1) if a person is black, then he or she will be a Democrat, and (2) if a person is white, then he or she will be a Republican. The proportion of times such predictions are correct for all cases is an index of the strength of the relationship. A high proportion indicates that the variables are strongly related; a low proportion indicates a weak relationship. Furthermore, if the proportion is so low that knowledge of one variable is of no use in predicting the other, the variables are said to be unrelated. The rest of Table 4.2 gives a possible combination showing a moderate relationship between race and political party affiliation (Table 4.2B) and a distribution indicating no relationship (Table 4.2C). Statistical indices of association may be computed for each of these distributions. Ordinarily, these indices will range from 0, indicating no relationship, to 1.00, indicating a perfect relationship.

Relationships among Quantitative Variables

As we move from qualitative to quantitative variables, it becomes possible to say whether a change in a variable represents an increase or a decrease in value. With this additional property, we can measure not only the strength of the relationship but also two other aspects: direction and linearity.

A relationship may be either positive or negative in direction. A **positive** or **direct relationship** between variables exists if an increase in the value of one variable is accompanied by an increase in the value of the other or if a decrease in the value of one variable is accompanied by a decrease in the value of the other. In other words, the two variables consistently change in the same direction. We would expect a positive relationship between sons' heights and fathers' heights (the taller a father, the taller his son will tend to be) and between students' scores on the SAT and their college grade-point averages (as scores increase, grades tend to increase). A **negative** or **inverse relationship** between variables exists if a decrease in the value of one variable is accompanied by an increase in the value of the other. Thus, changes in one variable are opposite in direction to changes in the other. We would find a negative relationship between a person's age and how long he or she is expected to live (as age increases, life expectancy decreases) and between the speed and accuracy with which people perform many tasks (the faster one does something, the less accurately one is likely to do it).

Relationships among quantitative variables are usually depicted with graphs, such as those in Figure 4.2. The lines in these two graphs illustrate the characteristic of "linearity." The straight line in Figure 4.2A depicts a *linear* relationship. (Specifically, this line depicts a *positive linear* relationship. A straight line running from the upper left corner to the lower right corner would depict a *negative linear* relationship.) Note that one variable changes at the same rate and in the same direction (positive) over the entire range of the other variable. The curved line in Figure 4.2B depicts a *curvilinear* relationship. In this case, the rate of change in one variable is *not* consistent over all values of the other: Variable Y increases more rapidly for low values than for high values of variable X and then reverses direction. We would expect this pattern of relationship to occur between the age and annual earnings of adult workers. Earnings will generally increase with age up to retirement and then will decline.

Tabular and graphic representations of relationships between variables and other statistical analyses are discussed further in Chapter 15. There, as well as elsewhere in the book, we focus on linear relationships since this pattern is most often assumed and analyzed in the social sciences. A common statistical measure of the strength and direction of linear relationships between two quantitative variables is the Pearson product-moment coefficient of correlation, or **correlation coefficient** for short. Symbolized by the letter r, the correlation coefficient may vary between -1.00 and $+1.00$. The signs $-$ and $+$ indicate the direction of the relationship, positive or negative; the magnitude of the coefficient, ignoring the sign, indicates the strength of association. Again, strength of association can be construed as accuracy of prediction. Relationships between two quantitative variables are represented in graphs such as Figure 4.2; however, when each case is plotted according to its values on the independent and dependent variables, the plots seldom form a perfectly straight line, as in Figure 4.2A. Rather, the plots cluster to varying degrees about the best-fitting line. If we use this line to predict Y values from X values, then the more closely the plots cluster about the line, the more accurate the prediction and the stronger the association. Figure 4.3 shows two possible outcomes. In Figure 4.3A, where the plots cluster fairly tightly around the line, there is a relatively strong correlation of .74 between variable X and variable Y. This is the

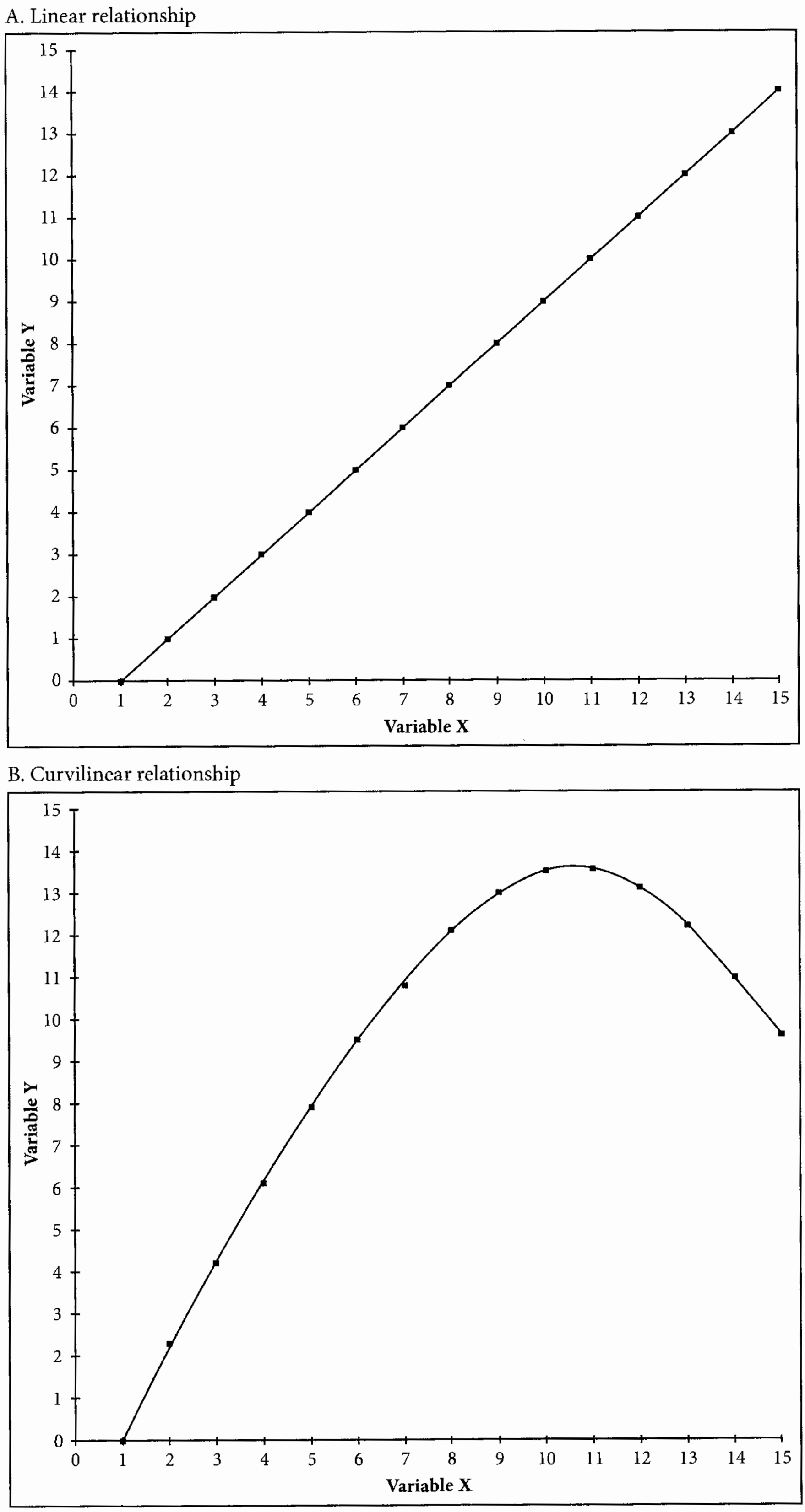
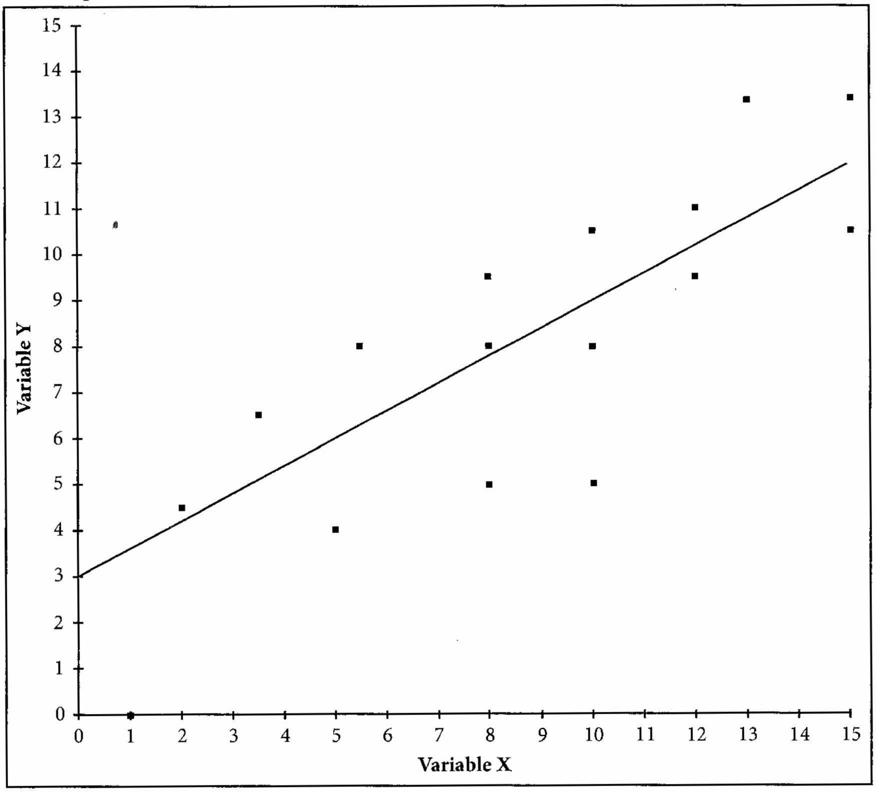


FIGURE 4.2. **Linear and curvilinear relationships between two quantitative variables.**

A. Strong correlation, $r = .74$



B. Weak correlation, $r = .08$

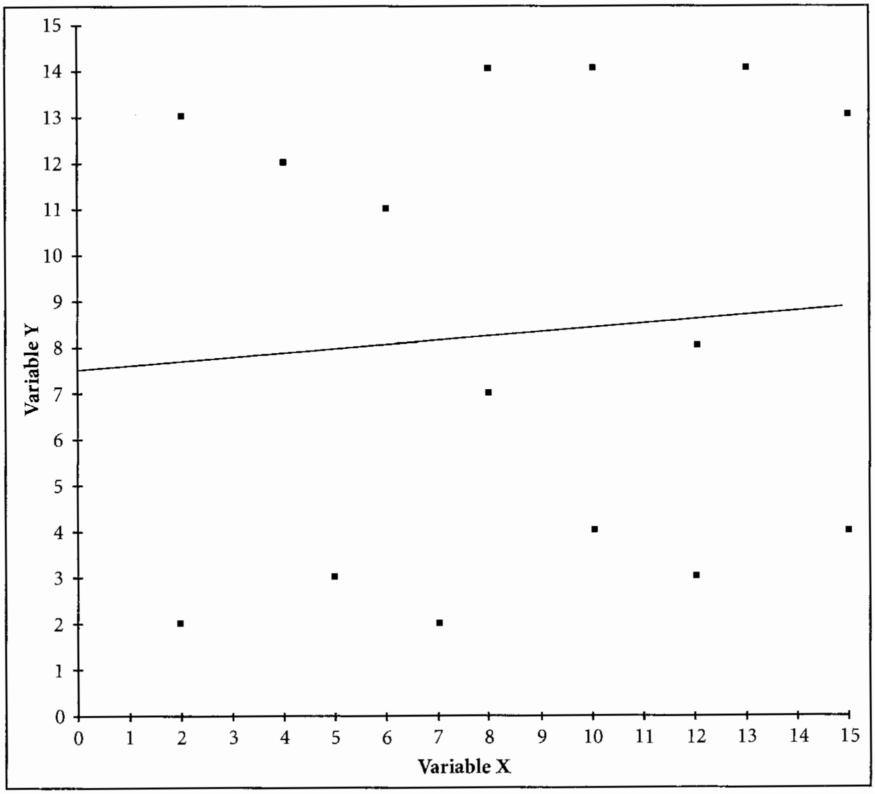


FIGURE 4.3. Strong and weak correlations between two quantitative variables.

same correlation that Broh (2002) found between students' grades in English and grades in math. In Figure 4.3B, where the plots are loosely clustered about the line, the correlation is .08. This is the correlation Broh found between number of hours spent on homework and grades in English.

Relationships between a Qualitative and a Quantitative Variable

Another method of assessing relationships is used when both qualitative and quantitative variables are involved. Most often in such cases, especially in experiments, the independent variable is qualitative and the dependent variable is quantitative. This is the type of relationship considered here.

It will be helpful, once again, to think in terms of predictions. A relationship is said to exist if the different categories of the independent variable predict different values for the dependent variable. Thus, if each category of the independent variable is treated as a distinct group, then a relationship can be described in terms of the *differences among groups* on the dependent variable. We might compute an average value of the dependent variable for each group. No differences in these averages across groups would then indicate no relationship. And, in general, the larger the differences, the stronger the relationship.

For example, suppose a researcher was interested in examining the relationship between race and annual income. Since annual income, as measured in dollars earned, is a quantitative variable, average incomes could be computed for each racial group (i.e., for each category of the qualitative variable race), as shown in Table 4.3.[3] Note that there is a difference in the averages across groups: In 2014, whites earned $56,866 on average and blacks earned $35,398. Thus, we may conclude that a relationship exists between race and income.[4]

Statistically Significant Relationships

Social scientists do not ordinarily consider an association between two variables meaningful unless the association is **statistically significant**. Consider the moderate association between race and party affiliation in Table 4.2B. Note that there are only 20 cases in this table. Now suppose that these 20 cases represent a random sample of the U.S. population. It would not be appropriate to conclude from these data that there is an association between race and party affiliation in the general population. Why? Because in a sample this small the difference in party affiliation between blacks and whites easily could have occurred by chance (i.e., as a result of random selection) even if there is no relationship in the larger population.

TABLE 4.3. Relationship between a Quantitative Variable (Median Income in 2014 Dollars) and a Qualitative Variable (Race)

	White	*Black*
Median income	$56,866	$35,398

Source: C. DeNavas-Walt and B. D. Proctor. *Income and Poverty in the United States: 2014*. U.S. Census Bureau, Current Population Reports, P60-252. Washington, D.C.: U.S. Government Printing Office, 2015, Table 1.

TABLE 4.4. Race by Political Party Preference,* 2014 General Social Survey

	White	*Black*
Democrat	49.9%	93.6%
Republican	50.1	6.4
	100.0%	100.0%
(*N*)	(994)	(266)

*Party preference indicates respondents who described themselves as either "strong" or "not very strong" Democrats (Republicans); data exclude those who identified as "independent." Based on 2196 white and black respondents.

Statistically significant associations are those that are *not* likely to have occurred by chance or random processes. Table 4.4 shows the relationship between race and party affiliation in a recent national sample of the adult population. The respondents in this survey were asked whether they usually thought of themselves as a Democrat, Republican, or Independent. Table 4.4 shows only those respondents who described themselves as either Democrats or Republicans. Blacks were far more likely than whites to identify themselves as Democrats (93.6 percent compared to 49.9 percent); given the size of the sample (2196), this difference is highly unlikely to have occurred by chance and therefore is statistically significant. When this property of relationships is reported in the research literature, it is indicated with an italic lowercase p followed by < and a decimal; for example, $p < .01$. This means that the probability is less than 1 in 100 that the relationship could occur by chance, assuming that there is no relationship in the larger population from which the sample is drawn. With odds this low, we would conclude that the association did not occur by chance and therefore that a relationship between race and party affiliation exists in the United States. (See Chapter 15 for a brief discussion of tests of statistical significance.)

It is important to note that statistical significance is not the same as strength of association. Although strong associations are more likely to be statistically significant than weak ones, a very weak association may be significant provided that the sample is large enough. For example, both of the correlations reported above for the Broh (2002) study (see Figure 4.3) were statistically significant, although one (.74) was strong and the other (.08) was very weak. The reason that such a low correlation was significant was that it was based on a sample size of more than 10,000 students. With samples this large, even relatively weak associations are unlikely to occur by chance and, therefore, are likely to differ from a zero association in the population (all high school students) from which the sample was drawn.

So far we have examined some common ways of depicting relationships. Also, we have noted four properties of relationships: strength, directionality, linearity, and statistical significance. The *strength* of a relationship refers to the extent to which variables are associated or correlated. *Directionality* and *linearity* tell us how changes in one variable are related to changes in another. Do the variables change in the same or the opposite direction? Is the rate of change in one variable consistent over all values of the other? Together these two properties describe the *form* of the relationship. Knowing the strength and form of relationships as well as whether it is statistically significant will often satisfy a researcher's curiosity.

However, if our interest is in explaining social phenomena, we also will need to know about the causal link between variables. And these statistical properties are never sufficient to establish a cause-and-effect relationship.

KEY POINT

Relationships may be described by their *form* (how changes in one variable vary with changes in the other), *strength* (how accurately values of one variable predict values of the other), and *statistical significance* (likelihood that the relationship occurred by chance or random processes).

The Nature of Causal Relationships

In Chapter 2 we noted that, for purposes of explanation and prediction, scientists find it helpful to think in terms of cause-and-effect relationships. But how does one identify such a relationship? What is meant by the term "causality"? At first glance, the task of defining causality seems simple. Since we are so accustomed to thinking causally and causal terms are so frequently used in everyday life, the concept would appear to be widely understood. In lay terms, a *cause* is something that makes something happen or change. It seems obvious that a rock thrown against a window will cause the glass to shatter. And the fact that drinking too much soda causes someone to get a stomachache is a **causal relationship** that you can comprehend even if you fortunately have not had the same experience.

In contrast to the implicit understanding of causality that seems to exist in everyday life, the meaning of the concept of "cause" has been hotly debated by philosophers and scientists for centuries. Much of this debate stems from the philosopher David Hume's analysis (1748). Hume argued that all that one can observe is a constant or stable association between events. From such association we infer a causal connection; however, there is no way of logically or empirically showing that a causal connection actually exists. A causal relationship exists only in the observer's mind; it is something inferred from an observed association between events. Following Hume's line of reasoning, some philosophers and scientists have argued against the use of the concept of cause in science. Yet many scholars regard causal relationships as the heart of scientific understanding, at least in the social, behavioral, and biological sciences (Woodward, 2016). Furthermore, even if such relationships cannot be "proven" empirically (just as no generalization can be proven by scientific evidence), researchers have found working with causal hypotheses to be a productive way of doing science (Blalock, 1964:6).

The important point of Hume's analysis, therefore, is that we should understand the bases for making causal inferences. In other words, what kind of evidence supports the belief that a causal relationship exists? Social scientists generally require at least three kinds of evidence to establish causality. These requisites are association, direction of influence, and nonspuriousness.

Association

For one variable to be a cause of another, the variables must have a statistical **association**. If the pattern of changes in one variable is not related to changes in another, then the one cannot be producing, or causing, changes in the other. Thus, for instance, if intelligence is unrelated to delinquency—that is, if adolescents of high and low intelligence are equally likely to be delinquent—then intelligence cannot be a cause of delinquency.

Associations are almost never perfect, so a perfect association between variables is not a criterion of causality. According to logicians, in fact, the very idea of causation implies imperfect associations between events. Causes can have invariable effects only in "closed systems" that exclude all other factors that might influence the relationship under investigation. Many of the laws of physics, for instance, are said to apply exactly only in a vacuum. However, vacuums are not found in nature; neither is it possible in real social situations to eliminate completely the influence of extraneous factors. Perfect associations may be expected, therefore, only under the theoretical condition that "all other things are equal" but not in the "real world" of observations.

Barring "perfect" associations, then, the application of this first criterion necessarily involves a judgment about whether an association implies a meaningful causal relationship. In the social sciences, causal relationships often are implied from comparatively "weak" associations. One reason for this is that many measurements in the social sciences are relatively imprecise. The primary reason, though, is that in explaining human action, multiple causes may independently or jointly produce the same or similar effects. A weak association may mean that only one of several causes has been identified, or it may mean that a causal relationship exists but not under the conditions or for the particular segment of the population in which the weak association was observed. Rather than strength of association, therefore, social scientists rely on tests of statistical significance to determine whether a meaningful, interpretable association exists between variables. Although Broh (2002) found a weak correlation between time spent on homework and grades, because this association was statistically significant, it was interpreted as a cause (among others) of academic performance.

Direction of Influence

A second criterion needed to establish causality is that a cause must precede its effect, or at least the **direction of influence** should be from cause to effect. In other words, changes in the causal factor, or independent variable, must influence changes in the effect, or dependent variable, but not vice versa. For many relationships in social research the direction of influence between variables can be conceived in only one way. For example, characteristics fixed at birth, such as a person's race and gender, come before characteristics developed later in life, such as a person's education or political party affiliation; and it is hard to imagine how changes in the latter could influence changes in the former.

Direction of influence is not always so easy to determine. Suppose you found a correlation between racial prejudice and interracial contact showing that the more contact a person has with members of other races, the less prejudiced he or she is apt to be. One possible interpretation is that racial contact increases familiarity and contradicts stereotypes, thereby reducing prejudice. An equally plausible interpretation is that prejudiced people will avoid contact, whereas tolerant people will readily interact with other races, so that racial prejudice influences racial contact. Without any information about the direction of influence between these variables, there is no basis for deciding which variable is the cause and which is the effect. To take another example, a correlation between grades and class attendance may mean that greater attendance "increases the amount learned and thus causes higher grades" or it may mean that "good grades lead students who obtain them to attend class more frequently" (Neale and Liebert, 1986:89–90).

Direction of influence was an issue in the Broh (2002) study. An association between sports participation and academic performance could mean that playing sports has educational benefits, but it also could mean that higher-achieving, "good" students are more likely to choose or be selected to play sports than other students. Because Broh hypothesized that sports participation had a positive influence on grades, it was important for her to rule out the possibility that superior academic performance leads to sports participation.

The requirement that causes influence effects has two major implications for research: (1) hypothesized relationships should always specify the direction of influence among variables and (2) whenever the direction cannot be established theoretically, it should be tested empirically. As you will see, this task is relatively easy in experiments but often difficult in other kinds of research.

Nonspuriousness (Elimination of Rival Hypotheses)

If two variables happen to be related to a common extraneous variable, then a statistical association can exist even if there is no inherent link between the variables. Therefore, to infer a causal relationship from an observed correlation, there should be good reason to believe that there are no "hidden" factors that could have created an accidental or spurious relationship between the variables. When an association or correlation between variables cannot be explained by an extraneous variable, the relationship is said to be *nonspurious*. When a correlation has been produced by an extraneous third factor and neither of the variables involved in the correlation has influenced the other, it is called a **spurious relationship**.

The idea of spuriousness is obvious when we consider two popular examples in the social sciences. The first is a reported positive correlation in Europe between the number of storks in an area and the number of births in that area (Wallis and Roberts, 1956:79). This correlation might explain how the legend that storks bring babies got started, but it hardly warrants a causal inference. Rather, the correlation is produced by the size of the population. Storks like to nest in the crannies and chimneys of buildings; so as the population and thus the number of buildings increases, the number of places for storks to nest increases. And as the population increases, so does the number of babies. We also would expect to find a positive correlation between the number of firefighters at a fire and the amount of damage done. But this does not imply that firefighters did the damage. The reason for the correlation is the size of the fire: Bigger fires cause more firefighters to be summoned and cause more damage.

In these examples the original relationship is an incidental consequence of a common cause: an antecedent extraneous variable. In the first example, the size of the population accounts for *both* the number of storks and the number of births; in the second, the severity of the fire determines *both* the number of firefighters and the amount of damage. These relationships are depicted in Figure 4.4. The examples are intuitively obvious, and the third factor is fairly easy to identify.

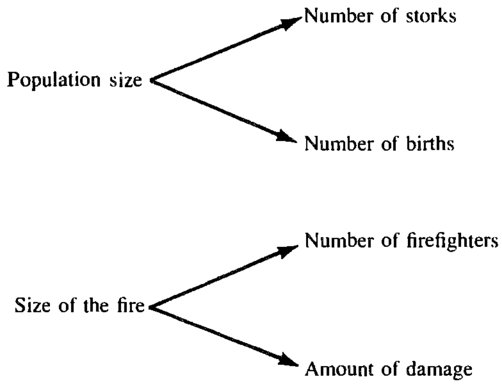


FIGURE 4.4. **Examples of spurious relationships.**

In actual research, spurious relationships are much less apparent, and the possibility often exists that an unknown variable may have produced an observed association. For many years, numerous studies have shown that children who are breastfed tend to have higher IQ scores than those who are not. Proponents of breastfeeding (which does have many other benefits for mother and child) have inferred a causal connection, contending that the effect may be due to a component of breast milk or perhaps to the physical interaction between mother and baby. Recent research suggests, however, that this association is spurious. Both breastfeeding and child intelligence are influenced by the mother's intelligence: Mothers with high IQs are more likely to breastfeed *and* more likely to have children with high IQs (Der, Batty, and Deary, 2006).

To infer that a relationship is nonspurious, researchers must identify and control for extraneous variables that might account for an association. We can see how this works by considering the relationship between cigarette smoking and lung cancer. Based on studies of U.S. men and women over age 55, the Centers for Disease Control and Prevention (2015) reported that those who smoked cigarettes were 25 times more likely to get lung cancer than people who did not smoke. But the reason scientists now believe that there is a causal link between smoking and cancer is not based simply on this association; rather, it is based on the fact that smokers' greater risk of getting cancer remains about the same when the effects of other variables are removed.

Suppose we control for the variable urban–rural residence. This could create a spurious association between smoking and cancer if urban areas have more smokers as well as sources of lung cancer (e.g., greater air pollution) than rural areas. To control for extraneous variables we must remove their effects, such as by computing the incidence of lung cancer among smokers and nonsmokers separately for urban and rural residents. A finding of no difference in the cancer rate between smoking and nonsmoking urban dwellers and between smoking and nonsmoking rural residents would suggest that urban–rural residence produced a spurious association between smoking and cancer. If we found, however, that both urban and rural smokers were more likely to develop lung cancer than nonsmokers in these areas, we would be more confident that the relationship is nonspurious.

To infer nonspuriousness, researchers must show that the relationship is maintained when all extraneous variables are held constant. Circumstances seldom allow for the control of all variables; therefore, researchers attempt to control the effects of as many as possible. The greater the number of variables that are controlled without altering a relationship, the greater the likelihood that the relationship is not spurious. Thus, we would become even more confident in a causal link between smoking and lung cancer if the incidence of lung cancer among smokers and nonsmokers remains the same when examined separately for men and women, people of low and high socioeconomic status, people of different races, and so forth. (Box 4.1 provides another example of the importance of the nonspuriousness criterion. In this case, the causal interpretation that exercise reduces the risk of heart attacks was challenged by several rival explanations.)

KEY POINT

An association between variables does not necessarily imply that the variables are causally related. In addition to association, causal inferences require evidence of direction of influence and nonspuriousness.

In one way or another, tests for spuriousness entail controlling for extraneous variables. The type of statistical control employed in our smoking–cancer example is common in nonexperimental research. Its major drawback is that one can control statistically only for those variables that have been observed or measured as part of the research. Hence, the effects of any unknown or unmeasured variables cannot be assessed. A stronger test of nonspuriousness is provided in experiments through a process called "randomization" that makes it theoretically possible to control for all extraneous variables. Experimental controls are discussed in Chapter 7, and causal analysis techniques involving statistical manipulation of nonexperimental data are discussed in Chapter 16.

BOX 4.1 Problems in Causal Interpretation: The Case of Exercise and Heart Attacks

Of the three criteria needed to establish a causal relationship, the most difficult to assess is nonspuriousness. One can never be sure that a causal connection exists between correlated variables. Indeed, mistaken impressions of causality may remain undetected for years. An interesting example of this problem in social science research is related by the psychologists Schuyler Huck and Howard Sandler (1979:151, 152, 227).

In recent years there has been much interest in the relative benefits of regular exercise. One controversial claim is that exercise can reduce the risk of heart attacks. An early study by J. N. Morris of London shows, however, just how difficult this is to establish. Examining drivers and conductors of London's double-decker buses, Morris found that the drivers were far more likely to suffer from heart disease and to die from coronaries than were the conductors. Since the drivers sat in their seats all day while the conductors ran up and down stairs to collect fares, he concluded that it was the differential amount of exercise inherent in the two jobs that brought about the observed differences in health. Before reading further, you might try to think of variables other than exercise that could have produced the difference in heart problems between the drivers and conductors. Morris uncovered one variable in a follow-up study, and Huck and Sandler mention two others.

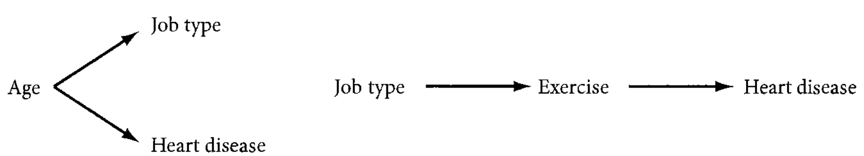
Some time after the publication of the above results, Morris examined the records maintained on the uniforms issued to drivers and conductors and discovered that drivers tended to be given larger uniforms than conductors. Therefore, he concluded, differences in weight rather than exercise might be the causal factor. That is, heavier men, who were more coronary-prone to begin with, may have chosen the sedentary job of driver, whereas thinner men chose the more physically active job of conductor.

Another explanation is related to the amount of tension associated with the two jobs. As Huck and Sandler (1979:227) point out,

> The conductors probably experienced very little tension as they went up and down the bus collecting fares from the passengers; the worst thing that they probably had to deal with in their jobs was a passenger who attempted to ride free by sneaking around from one seat to another. Normally, however, we suspect that the conductors actually enjoyed their interaction with other people while on the job.
>
> But on the other hand, each driver had the safety of everyone on the bus as his responsibility. And as anyone who lives in or visits a city knows, driving in rush-hour traffic is anything but restful. Having to dodge pedestrians, being cut off by other vehicles, watching for signal changes—these activities can bring about temporary outbursts of anger and chronic nervousness. Imagine how it would affect your heart to be in the driver's seat of a bus for eight hours each working day!

Finally, a third variable that could account for the different rate of heart problems is age. If mobility or seniority or some other function of age were related to job assignment, then employees assigned to the driver jobs may have been older and those assigned to the conductor jobs younger. And since we would expect more heart attacks among older persons, age rather than the nature of the job could be the causal variable.

In this example it is impossible to tell which cause—exercise, weight, job stress, or age—may have produced the observed differences in health between drivers and conductors. Since both weight and age are antecedent to job type and heart disease, either of these uncontrolled extraneous variables could have created a spurious relationship. However, if exercise or job stress were the correct interpretation, then the original relationship would not be spurious, since exercise and job stress specify intervening variables through which the job itself can make a person more or less susceptible to heart problems. The following diagram shows the difference in these two outcomes.



(continued)

(continued)

It is possible that two or more of these variables are operating jointly to produce the health differential between the two groups. The only safe conclusion is that we really do not know which interpretation is correct.

In general, correlation does not imply causation. All correlations must be interpreted; like any fact, they do not speak for themselves. To infer a causal relationship from a correlation, an investigator must detect and control for extraneous variables that are possible and plausible causes of the variable to be explained. The fatal flaw in Morris's study is that relevant extraneous variables were not controlled; without directly assessing the effects of such "hidden" causes, we cannot tell which interpretation is valid.

Causation, Intervening Variables, and Theory

Recall that extraneous variables may be either antecedent or intervening in relation to the independent and dependent variables. An antecedent variable, which is causally prior to both the independent and the dependent variable, can produce a spurious relationship. By contrast, the identification of an intervening variable or mechanism linking the independent and dependent variables strengthens the causal inference. Indeed, this is sometimes advocated as a fourth criterion—in addition to association, direction of influence, and nonspuriousness—for establishing that one variable causes another (see Hyman, 1955). For example, one may argue that the belief that smoking causes lung cancer will be enhanced considerably if and when it is established that certain chemical agents from cigarettes produce cancerous cells. Knowing the causal process through which smoking produces cancer would provide one last shred of evidence against a spurious correlation. However, once the criterion of nonspuriousness is firmly established, a causal relationship is generally inferred even if the intervening mechanisms are not known. Few scientists today doubt that smoking causes cancer. Thus, although specifying the intervening variables in a relationship may lead to a better theoretical understanding and more accurate prediction, it is "not part of the minimum requirements for demonstrating causality. Holding a match to a pile of leaves is a cause of their bursting into flame, even if one cannot describe the intervening chemical reactions" (Hirschi and Selvin, 1967:38).

KEY POINT

Antecedent variables may create spurious associations; *intervening variables* clarify the causal connection between variables.

Although not a necessary causal criterion, the identification of intervening variables is nonetheless an essential part of scientific inquiry. Often, in fact, this is what the development and testing of theory is all about. For example, Durkheim's theory of suicide stipulated the intervening causal mechanism—social integration—for a number of relationships. The reason that fewer suicides are found among Catholics than among Protestants and that there are fewer suicides among married than among single people, according to the theory, is that being a Catholic and being married each engenders a greater sense of social integration, which in turn reduces the likelihood that anyone within the group will commit suicide. Broh (2002) tested the effects of several intervening variables. Having found that participation in interscholastic sports benefits students' academic performance, she also showed that this relationship was best explained by developmental and social capital theories. High school athletes develop a greater work ethic and sense of control over their lives than do nonathletes, and their athletic participation generates stronger ties among students, parents, and teachers, all of which have positive effects on academic performance.

Yet, theory plays a much larger role in causal analysis than specifying intervening variables. Theories not only render a more complete understanding of the causal processes that connect events but also provide the general framework for investigating the nature of all relationships. Theories tell a researcher which relationships to observe, what extraneous variables are likely to affect the relationships, and the conditions under which a causal relationship is likely to exist. It is only in terms of some theory, in short, that a researcher can determine how to assess the meaningfulness of a weak association and how to test for direction of influence and nonspuriousness. Thus, we see again the importance of the interplay between theory and research in science. Theory guides research, and research provides the findings that validate and suggest modifications in theory.

Formulating Questions and Hypotheses

Having introduced the language of units of analysis, variables, and relationships, we are now ready to return to formulations of research questions. According to Glenn Firebaugh (2008), research questions should meet two fundamental criteria: They should be researchable and interesting. A researchable question is one that is answerable through empirical research. As we saw in the preceding chapters, some questions cannot be answered because they are beyond the realm of science (e.g., Does God exist? Is capital punishment morally wrong?) or because ethical considerations rule them out. Other questions may be too grand. Novice as well as experienced researchers often begin with questions that are so broad in scope that they provide little or no immediate direction for research. For example, Stanley Schachter, well known in part for his work on affiliation, was interested in what motivates people to be around others (Evans, 1976:159), a question nearly as encompassing as the entire field of social psychology. Before Schachter could begin to do research, he had to reformulate this question so that it pointed to identifiable variables and relationships. After reading and speculating about "affiliative tendencies," Schachter eventually focused on the question of how fear affects the desire to be with others.

Firebaugh's second criterion for good research questions is that they should be interesting to you and, especially, to others. What makes a question interesting? Primarily it should contribute to ongoing conversations in a scientific field of study. The only way to know whether your question speaks to such conversations is to learn about the latest research through a thorough review of the literature. Here is where you will find the major theories, key findings, and unresolved issues in the field. Beckett Broh's (2002) literature review revealed several studies that found a positive association between high school students' participation in sports and grades. Relatively few studies, however, had examined the effect of other extracurricular activities, and even fewer had attempted to test theoretical explanations of how sports participation enhances educational achievement. After considering various theories, Broh formulated three unanswered research questions:

- Why does sports participation boost students' achievement? Does sports participation benefit students' development and social networks, and are these the mechanisms that link participation to educational outcomes?
- Are the educational benefits of sports participation unique to sports, or do nonsports extracurricular activities also promote achievement?
- Do nonsports extracurricular activities benefit students' development and social networks? (Broh, 2002:73)

Broh's research questions were "interesting" because they extended prior research in new directions. (Also of interest, albeit much less common, are research questions that challenge well-established findings.) Her questions extended research in two ways: by broadening the question to include participating in nonsports activities as well as interscholastic sports and by identifying an intervening variable (social capital) that accounts for the effect of sports participation on grades. Two other ways of extending prior research suggested by Firebaugh (2008) are investigating whether a finding applies to a new population (imagine Broh replicating her study outside the United States) and whether a finding applies to a new time period (e.g., would a similar association be found in the first half of the twentieth century?).

The tentative answers to research questions are called "hypotheses." Formally defined, a **hypothesis** is an expected but unconfirmed relationship between two or more variables. Hypotheses come from a variety of sources, including everything from theory to direct observation to guesses and intuition. Sometimes the formulation of hypotheses is the principal outcome of research. At other times, hypotheses are never made explicit, even though they implicitly guide research activities. However, whenever the research objective is clearly one of testing relationships among variables, hypotheses should be stated formally and precisely so that they carry clear implications for testing the stated relations.

Although they may be stated in a variety of ways, all hypotheses should speculate about the nature and form of a relationship. An adequate hypothesis statement about two variables should indicate which variable predicts or causes the other and how changes in one variable are related to changes in the other. For example, if we thought that education generally increases tolerance, we might hypothesize that "an increase in education will result in a decrease in prejudice." This statement implies two features about the relationship: first, which variable causes, explains, or predicts the other (education predicts prejudice); and second, how changes in one variable are related to changes in the other (as education increases, prejudice decreases). Hypotheses that specify the form of the relationship are said to be testable because it is possible, assuming each variable has been measured adequately, to determine whether they are true or false or at least whether they are probably true or probably false.

There are several ways of expressing testable hypotheses, such as the one relating education and prejudice:

1. *"If–then" (conditional) statements.* These statements say that if one phenomenon or condition holds, then another will also hold. An example is "If a person has a high level of education, then he or she will have a low level of prejudice." Alternately, one could say, "If a person has a low level of education, then he or she will have a high level of prejudice." In logic, such statements are called **conditionals**. A conditional consists of a connection between two simple statements, each pointing to a condition or category of a variable. As used in science, the connection asserted by a conditional is that the condition (or variable category) following the "if" *causes* the condition (or variable category) following the "then." Social research hypotheses seldom are specified in this form. Still, it is always possible to restate testable hypotheses as conditionals; and using this standard logical form, some scientists contend that it is easier to ascertain the kinds of inferences that legitimately can be made from research findings to the hypothesis (McGuigan, 1993:37).
2. *Mathematical statements.* Some hypotheses may be stated in the form of the standard mathematical formula $Y = f(X)$, which reads "Y is a function of X." An example is Einstein's famous formula $E = mc^2$ (i.e., energy equals mass times the speed of light squared). Although it was once a hypothesis, $E = mc^2$ is now called a scientific "law" because it has been confirmed repeatedly. Mathematical formulas represent precise formal statements of hypotheses. Because variables in social research generally are measured with less precision than in the physical sciences, social scientists seldom state hypotheses in this form. Nonetheless, mathematical formulas are the ideal in science because they yield precise predictions and express complex relationships parsimoniously. The mathematical form is also equivalent to the conditional. $Y = f(X)$ merely says "If (and only if) X is this value, then Y is that value" (McGuigan, 1993:38).
3. *Continuous statements.* Hypotheses of the form "The greater the X, the greater (or lesser) the Y" indicate that increases in one variable (X) are associated with increases (or decreases) in another variable (Y). For example, as education increases, prejudice decreases; or, expressed in slightly different form, the higher the level of education, the lower the prejudice.
4. *Difference statements.* Statements in this form assert that one variable differs in terms of the categories of another variable. For example, people with high education are less prejudiced than people with low education. Whether "continuous" or "difference" statements are used to express a hypothesis will depend on whether the variables in the hypothesis are quantitative or qualitative. If both variables could be "quantified," then the relationship could be stated in the continuous form. But if either variable consisted of discrete categories, such as "high" and "low" prejudice, then the relationship would need to be stated in the difference form. In either case, it is easy to transform the hypothesis statement into a conditional. For example, the statement "The higher the level of education, the lower the level of prejudice" becomes "If education is high, then prejudice will be low."

Both continuous and difference statements clearly specify the *form* of the relationship. However, both types of statements are ambiguous about the causal connection between variables. Ordinarily, a statement of the form "The greater the *X*, the greater the *Y*" is meant to imply that *X* causes *Y*. But it also can mean that *Y* causes *X* or that *X* and *Y* cause one another. As noted in our discussion of causation, it is important to know which variable is presumed to cause the other; in addition, since most hypotheses take the form of continuous or difference statements, the causal linkage can be problematic. Fortunately, in most research articles and reports, researchers make this connection clear in their discussions of the hypothesis.

At this point we should reiterate an important tenet of scientific research. Although all of the above forms of expression appear to assert that the relationship is absolutely true or false, hypotheses in science can only have probabilistic, not exact, confirmation. Thus, although the statement "If *X*, then *Y*" logically can only be true or false, it is assumed that observations will show it to be "probably true" or "probably false."[5] One often sees this kind of assumption built into statements of hypotheses in the form of qualifiers such as "tends to" or "in general." For example, one might say that "increased education will *tend* to reduce prejudice." Such statements acknowledge that tests of hypotheses are always restricted by the limited accuracy of measures and the inability to specify and control all the variables affecting events.

The idea that hypotheses can only be judged to be *probably* true or *probably* false gives rise to another type of hypothesis associated with tests of statistical significance. Significance tests indicate the probability that an observed relationship between two variables is due to chance variation; if a test is significant, this means that there is a low probability, say less than 5 in 100, that the relationship occurred by chance, and we therefore conclude that the relationship exists or is real. In statistics, the hypothesis that a relationship is a chance occurrence, which is the opposite of the hypothesized relationship, is called the **null hypothesis**. In other words, evidence supports the research hypothesis only insofar as the null hypothesis is rejected with a known probability of error. Researchers do not always make reference to the null hypothesis when testing relationships; however, it is central to the logic of significance testing and is always implicit.

How a hypothesis is expressed in a given study will depend on several factors: the researcher's discretion, the current state of knowledge about the research problem, and whether qualitative or quantitative variables are involved. Regardless of how hypotheses are expressed, however, they should indicate at least the form of the relationship between variables and, ideally, should specify the causal linkage between variables; ultimately, it is causal relationships in which scientists are interested.

Research Purposes and Research Design

In the previous pages we have emphasized the role of relationships in social research. The sense that social scientists make of the social world is expressed in terms of relationships among variables. Furthermore, although they may be only vaguely defined or implicit and unknown to the researcher, anticipated relationships structure a researcher's every activity, from deciding which variables to measure to deciding how observations should be made and interpreted. Not all research is conducted for the immediate purpose of testing relationships, however. Research is undertaken for three broad purposes: (1) to *explore* a phenomenon such as a group or setting to become familiar with it and to gain insight and understanding about it, frequently to formulate more precise research questions for further study; (2) to *describe* a particular community, group, or situation as completely, precisely, and accurately as possible; and (3) to examine and formally *test relationships* among variables. Whether a study is conducted primarily for the purpose of exploration, description, or testing relationships is important to know because these three functions have different implications for research design.

Exploratory research is undertaken when relatively little is known about something, perhaps because of its "deviant" character or its newness. Falling into this category of research are observational studies of street gangs and radical political and religious movements; clinical case histories of persons, groups, and events; and anthropological accounts of entire cultures. When exploring a topic or phenomenon about which one knows very little, one necessarily begins with a general description of the phenomenon. This sounds easy but in fact is probably the most difficult kind of study for a novice researcher to undertake. There are no clearly delineated independent and dependent variables and, therefore, no preset categories of observation and analysis. A researcher may have few, if any, guidelines to help determine what is important, who to interview, or what leads to follow. For these reasons, the research plan in an exploratory study is more open than in other kinds of research. Decisions are made about the kinds of instruments needed (e.g., photographic equipment, tape recorders) and the key persons with whom one will need to speak at first, but the paths down which these initial steps may lead are almost impossible to foresee. In Chapter 11, on field research, we discuss some data-gathering approaches that entail exploration.

The objective of a **descriptive research**, as the name implies, is to describe some phenomenon. All scientific research involves description at some level. Exploratory field research, for example, usually attempts to describe in detail the setting or cultural group under study. A purely descriptive study, however, is much more structured. It is basically a fact-finding enterprise, in which information is gathered from a set of cases that are carefully selected to enable a researcher to make estimates of the precision and generalizability of the findings. Examples of descriptive studies are the various censuses conducted periodically at the local and national levels. A census may provide information about everything from the age and racial composition of a community to its employment and housing costs. Also having a descriptive purpose are the ubiquitous opinion polls that attempt to estimate the proportion of people in a specified population who hold certain opinions or views or who behave in certain ways. For example, how many favor capital punishment? How many say they will vote for candidate A in the next election?

The third purpose for which research is conducted is to test relationships. Research with this purpose is sometimes called **explanatory research** because it formally seeks to explain why phenomena occur by identifying causes and effects. Since all research involves description, the primary difference between descriptive and explanatory studies lies in the scope of the description. Purely descriptive research operates at a lower level of description by merely seeking information about isolated variables, whereas explanatory research goes beyond this step to a description of relationships among variables. A survey that simply reports attitudes toward gun control is descriptive, but one that examines why people favor or oppose gun control is explanatory.

Both descriptive and hypothesis-testing research are highly structured and must be carefully planned. With these kinds of research it is important, therefore, to have a complete, detailed strategy worked out before the data are collected. This preliminary strategy or outline is what we have called the **research design.** The research design consists of a clear statement of the research question as well as plans for gathering, processing, and interpreting the observations intended to answer the question. To formulate a research design is to anticipate the entire research process, from beginning to end. To do this, one must have an adequate knowledge of every stage of social research. Examining these stages now will serve not only to clarify the key components of research design but also to introduce the reader to the remainder of the book.

Stages of Social Research

Recall from Chapter 2 that the scientific process consists of the cyclical interplay between theory and data. At some point, scientists work deductively from theory to data in testing hypothesis; at another point, they interpret data and infer theories. Thus, scientific inquiry is alternately deductive and inductive. Figure 4.5 outlines the major steps that apply to deductive explanatory research, which is the primary focus of this book. Although most deductive inquiry follows this pattern, steps may be omitted, depending on the research

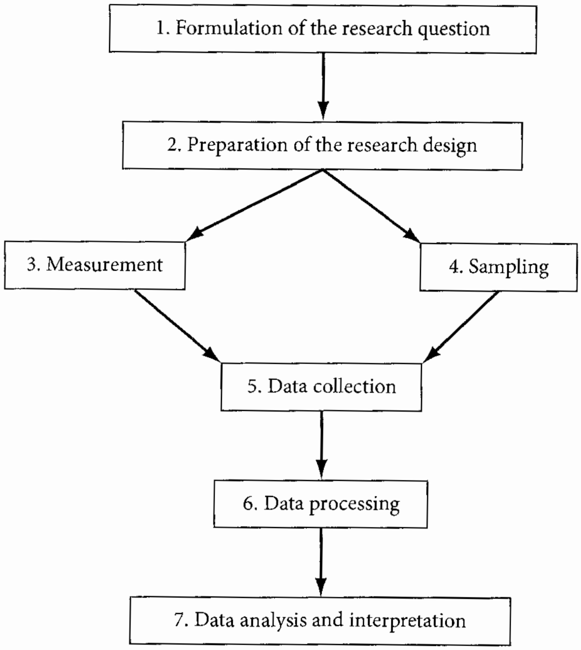


Figure 4.5. The stages of social research.

approach, and there may be numerous feedback loops. In short, this is an idealized model; some research is much less orderly than it implies.

Stage 1: Formulation of the Research Question

Research begins with a question. Questions initially chosen almost always require more precise formulation to be amenable to research. From a general idea, one must decide more specifically what one wants to know and for what purpose one wants to know it. The best ideas for formulating questions are likely to be found in the scientific literature. In addition to helping to narrow and refine the research question, a review of the literature may reveal its broader theoretical significance. Literature reviews also may help to identify relevant control variables and should suggest pertinent methods and procedures by indicating how other researchers have addressed the question.

Stage 2: Preparation of the Research Design

Once the question has been clearly formulated, a researcher must develop an overall plan or framework for the investigation. To do this, he or she must, in effect, anticipate all of the subsequent stages of the research project. Preliminary decisions must be made about what sort of observations are needed to answer the research question or to provide an adequate test of the hypothesis. The researcher must then select an appropriate strategy for making the observations—experiment, survey, field research, or use of available data. Each of these approaches has its unique strengths and weaknesses that determine its suitability for given questions. Often, the best strategy, as we argue in Chapter 13, is a combination of approaches. Within the context of selecting an overall strategy, decisions also must be made on the unit of analysis, on which variables to observe and control and how they should be measured, and on how best to analyze the data. Thinking through all of these issues in advance should prevent serious mistakes and omissions in a study. Not all problems can be foreseen, however, especially in exploratory research, and many of the decisions at the design stage will be arbitrary and subject to change.

Two issues—measuring variables and selecting units of analysis—warrant special attention and are ordinarily worked out in detail after the basic research design is complete. These are therefore considered separate, concurrent stages.

Stage 3: Measurement

Part of the research plan involves devising operations that will link specific concepts to empirically observable events. This process of operationalization and measurement as well as techniques for assessing the quality or goodness of measures will be taken up in Chapter 5.

Stage 4: Sampling

In addition to deciding on the unit of analysis, researchers must also determine how many units should be selected and how to go about choosing them. The problems and methods related to sampling are addressed in Chapter 6.

Stage 5: Data Collection

In designing a study, researchers first decide on a general approach to data collection, which affects decisions about measurement and sampling. As we have pointed out, there are four basic choices: experiment, survey, field research, and use of available data. The underlying logic of each of these approaches, as well as their distinctive issues, advantages, and disadvantages, is covered in Chapters 7 through 12.

Stage 6: Data Processing

Having made the observations, a researcher is ready to analyze and interpret them. Prior to analysis, the data must be transformed or processed. Chapter 15 begins with a discussion of data processing, giving special attention to the preparation of data for statistical analysis.

Stage 7: Data Analysis and Interpretation

Once the data are ready for analysis, they must be manipulated further so that their meaning and bearing on the questions and hypotheses that initiated the inquiry can be extracted. There are several types of analysis, many of which involve statistical tests that are beyond the scope of this textbook. Without any training in statistics, however, one can learn how to read and to present data properly in tables and graphs and when it is appropriate to apply certain statistical procedures. Also, it is possible to develop a solid understanding of the logic of analyzing causal relationships. These topics are also covered in Chapters 15 and 16. Having analyzed the data, one draws conclusions about the hypotheses and theory that guided the research and, if appropriate, assesses the practical implications of the findings. Finally, one reports the results.

Summary

- There are many sources of ideas and factors affecting topic selection, including theoretical and practical relevance, social premiums, and the personal values and resources of the researcher.
- General topics are refined and research questions emerge in the course of working out the overall study plan, or research design, the key elements of which are units of analysis, variables, and relationships.
- Units of analysis are the entities about whom or about which a researcher gathers information. These include individual people, groups and organizations of all kinds, communities, nations, and social artifacts.
- Researchers often aggregate information about individuals to describe the social unit that the individuals comprise. In general, conclusions should be restricted to the unit of analysis to which the data pertain. In drawing conclusions about individuals from group-level data, one risks committing an ecological fallacy.
- Variables are the characteristics of units that may vary in successive observations. Research focuses on explanatory variables while attempting to eliminate the influence of extraneous variables.
- The dependent variable is the explanatory variable that a researcher tries to explain or predict; an independent variable is a presumed cause of the dependent variable.

- Other variables may be either antecedent or intervening in relation to specific independent and dependent variables and either controlled or uncontrolled during the course of the study.
- Quantitative variables have categories that express numerical distinctions, whereas qualitative variables involve differences in kind rather than in number.
- Anticipated and predicted relationships guide all social research.
- Relationships occur when the changes in two or more variables form a predictable pattern. This pattern has two properties, strength and form, which are depicted differently according to whether the variables are qualitative or quantitative.
- In addition to statistical measures of strength of association, such as the correlation coefficient, tests of statistical significance indicate whether a relationship is likely to exist or whether it is likely to be the product of random processes.
- To establish that variable X *causes* variable Y, one must show that X and Y are statistically associated, that the direction of influence is from X to Y, and that the association between X and Y is nonspurious. Theories that identify intervening variables also may strengthen the inference that X causes Y.
- Research questions should be scientifically answerable and also interesting by contributing to ongoing research. Questions ask about a relation between variables, whereas a hypothesis is a conjecture about the nature and form of a relationship.
- In social research, a hypothesis may take the form of a conditional, a mathematical equation, a continuous statement, or a difference statement.
- Although much social research is undertaken to test hypotheses, it also may be done for the purposes of exploration and description.

Key Terms

unit of analysis
aggregate data
ecological fallacy
variable
    explanatory variable
    extraneous variable
    dependent variable
    independent variable
    antecedent variable
    intervening variable
    control variable
    quantitative variable
    qualitative variable
relationships among variables
    positive (direct) relationship
    negative (inverse) relationship
    correlation coefficient
    statistical significance
causal relationship
    association
    direction of influence
    spurious relationship
    nonspuriousness
hypothesis
conditional
null hypothesis
exploratory research
descriptive research
explanatory research
research design

Exercises

1. Suppose an anthropologist studies several villages in southern India. For each village she has data on the number of people and the average age of men and women, and she has computed an index of

the monetary value of various consumer goods in the village (e.g., the value of the total number of refrigerators, televisions, and air conditioners divided by the size of the population). An analysis of these data reveals that the villages with populations having the lowest average age also have the highest average dollar value of modern consumer goods (Bernard, 1994). What is the unit of analysis in this study? Can the anthropologist conclude from this association that young people purchase more modern consumer goods than older people? Why or why not?

2. For each of the following research questions or hypotheses, identify the *unit of analysis*, *independent variable*, and *dependent variable*.
   a. Boys whose parents are divorced or separated have more behavior problems than boys living with both biological parents.
   b. The number of sexually explicit magazines (such as *Hustler* and *Playboy*) sold in each of the 50 states is positively correlated with the number of reported rapes.
   c. Residents of rural communities are less tolerant of people holding controversial views than are urban residents.
   d. How does the level of economic development of a country affect the level of human services provided to its people?
3. The contact hypothesis, described recently as "one of the most durable ideas in the sociology of race and ethnic relations," has been investigated extensively over the past 50 years. In its simplest form, the hypothesis is that contact between members of different races fosters positive racial attitudes. Imagine you want to test this hypothesis. Specifically, you hypothesize that prior interracial contact in schools and neighborhoods will lead to greater racial diversity in one's current circle of friends. Assume you are going to do a survey at your college (or a local college).
   a. Suppose you find a correlation of .23 between prior interracial contact and racial diversity of friends. What does this indicate about the relationship?
   b. Now suppose the correlation was significant at $p < .01$. What does this information reveal about the relationship?
   c. An important control variable in your study would be race. Restate the hypothesis with race as a control variable.
4. In a campus survey, a researcher finds a statistically significant correlation of −.22 between frequency of exercise and depression; that is, the more students exercise, the less depressed they are. Can he conclude from these data that lack of exercise is one cause of depression? Why or why not? Now carefully explain how gender could create a spurious association between frequency of exercise and depression.
5. Based on a national survey, Christopher Ellison and Kristin Anderson (2001) examined the relationship between religious involvement and domestic violence among U.S. couples. Several previous studies had shown that regular religious attendance is inversely related to abuse among both men and women. The researchers analyzed several extraneous variables including respondents' age, education, employment status, and psychological depression.
   a. Select one of these variables and explain how it could create a *spurious association* between religious involvement and partner violence.
   b. The researchers hypothesized that "religious communities may reduce the risk of abuse by enhancing the levels of social integration and support enjoyed by their members." Are they hypothesizing that the relationship between religion and abuse is spurious? Explain.
6. Formulate three hypotheses, each relating two variables from the following list. For each hypothesis, specify the independent and dependent variables.

   Gender (male/female)
   Level of education (highest grade completed)
   Marital status (married/widowed/divorced/separated/never married)
   Party identification (Republican/Democrat/Independent/none)
   Belief in life after death (yes/no)
   Attitude toward capital punishment (favor/oppose)
7. Express one of the hypotheses formulated in Question 6 in the form of a (a) conditional, (b) continuous statement (if possible), and (c) difference statement.

Notes

1. When Frans van Poppel and Lincoln Day (1996) tested Durkheim's theory with individual-level data from the Netherlands at roughly the same period as Durkheim, they found that the Catholic–Protestant difference in suicide rates could be accounted for by how the causes of deaths were recorded. Catholic deaths were far more likely than Protestant deaths to be recorded as "sudden death" or "death from ill-defined or unspecified cause"; these deaths would have been categorized as suicides had they occurred among Protestants. Although Durkheim's theory may be supported by other data, this analysis shows the pitfalls of using group-level data to make inferences about individuals.
2. We caution readers not to confuse this distinction with the distinction between quantitative and qualitative research or quantitative and qualitative measurement. Both quantitative researchers (typically survey researchers and experimenters who rely on numbers and statistical methods) and qualitative researchers (who do intensive interviews, participant observation, and depth analyses of historical materials and rely on discursive methods) use both kinds of variables in their analyses (see King, Keohane, and Verba, 1994).
3. Table 4.3 reports a type of "average" called the "median," which is equal to the 50th percentile: Half of the group earn less and half earn more.
4. Determining the level of confidence that a relationship exists, as well as estimating the strength of the relationship, is the job of statistics and is therefore beyond the scope of this textbook. In general, confidence and strength will be greater to the extent that (a) the observed differences between groups are large, (b) the computed averages are based on a large number of people, and (c) people in the same group are similar to one another and different from those in other groups.
5. Determining whether a statement is probably true or probably false, like assessing the existence and strength of a relationship, is a matter of applying the laws of probability, which is the subject of statistics.