simplified. Evidence-based images emerge from the simplification of truth tables in the form of configurations of conditions that differentiate subsets of cases.

In many ways the comparative approach lies halfway between the qualitative approach and the quantitative approach. The qualitative approach seeks in-depth knowledge of a relatively small number of cases. When the focus is on commonalities, it often narrows its scope to smaller sets of cases as it seeks to clarify their similarities. The comparative approach usually addresses more cases because of its emphasis on diversity, and it is applied to sets of cases that are clearly bounded in time and space. As Chapter 6 shows, the quantitative study of covariation seeks broad familiarity with a large number of cases and most often views them as generic, interchangeable observations.

6

Using Quantitative Methods to Study Covariation

Introduction

The starting point of quantitative analysis is the idea that the best route to understanding basic patterns and relationships is to examine phenomena across many cases. Focusing on any single case or on a small number of cases might give a very distorted picture. Looking across many cases makes it possible to average out the peculiarities of individual cases and to construct a picture of social life that is purified of phenomena that are specific to any case or to a small group of cases. Only the general pattern remains.

Quantitative researchers construct images by showing the covariation between two or more features or attributes (variables) across many cases. Suppose a researcher were to demonstrate in a study of the top 500 corporations that those offering better retirement benefits tend to pay lower wages. The image that emerges is that corporations make trade-offs between retirement benefits and pay, with some corporations investing in long-term commitments to workers (retirement benefits) and some emphasizing short-term payoffs (wages and salaries). Evidence-based images such as these are general because they describe patterns across many cases and they are *parsimonious*—only a few attributes or variables are involved (pay and retirement benefits).

Images that are constructed from broad patterns of covariation are considered general because they condense evidence on many cases. The greater the number of cases, the more general the pattern. A quantitative researcher might construct a general image of political radicalism that links degree of radicalism to some other individual-level attribute, such as degree of insulation from popular culture, and use survey data on thousands of people (including people who are politically inert) to document the connection. Qualitative researchers studying this same question would go about the task very differently. The images they construct are detailed and specific, and they use methods that enhance rather than condense evidence. Using a qualitative approach, a researcher might construct an image of how political radicals nurture their radical commitments by studying the daily lives of twenty radicals in depth.

These two images of radicalism, one by a qualitative researcher and one by a quantitative researcher, might or might not contradict. Even if they did not contradict each other, the two images still would be very different in degree of detail and complexity. Quantitative researchers sacrifice in-depth knowledge of each individual case in order to achieve an understanding of broad patterns of covariation across many cases.

Quantitative researchers often use the term *correlation* to describe a pattern of covariation between two measurable variables. In the previous example, degree of radicalism and degree of insulation from popular culture are correlated such that more radical people tend to be more insulated. They also sometimes describe a correlation between two variables as a relationship, which should not be confused with the more conventional use of the term *relationship* to describe social bonds (for example, two lovers have a relationship). Again using the previous example, there is a relationship between degree of radicalism and degree of insulation.

Usually, attributes of cases that can be linked in this way are understood as variables because they are phenomena that vary by level or degree. There are cases with high values of a variable (for example, more than eighteen years of education on the variable "educational attainment"), cases with moderate values (say, twelve years of education), and cases with low values (only a few years of education). Some variables (called **independent** or causal variables) may be defined as causes, and others (called **dependent** or outcome variables) may be defined as effects in a given analysis. The dependent variable is the phenomenon the investigator wishes to explain; independent variables are the factors that are used to account for the variation in the dependent variable. A dependent variable in one analysis (for example, Gross National Product per capita in a study that seeks to explain why some countries are poor and others rich) may appear as an independent variable in the next (for example, as a causal variable that explains why people in some countries have a higher life expectancy than people in other countries).

The Goals of Quantitative Research

Because the quantitative approach favors general features across many cases, it is especially well suited for several of the basic goals of social research. These include the goals of identifying general patterns and relationships, testing theories, and making predictions. These three goals all dictate examination of many cases—the more, the better—and favor a dialogue of ideas and evidence that centers on how attributes of cases (variables) are linked to each other.

Identifying General Patterns and Relationships

One of the primary goals of social research is to identify general relationships. For a relationship to be general, it must be observed across many cases. In quantitative research this is understood not as observing the same exact phenomenon in each and every case, but as observing an association between two or more phenomena across many cases. When a social researcher claims that poorer countries tend to have higher rates of homicide, he or she in essence is stating that there is a general correspondence between a country's wealth and its rate of homicide such that richer countries tend to have lower homicide rates and poorer countries tend to have higher rates. (The United States is a striking exception to this general relationship.)

Identifying general patterns and relationships is important because they offer important clues about causation. It is obviously not true that if two variables are related across many cases, then one necessarily causes the other. If we found that shoe size and income were related, we would not argue that big feet cause high incomes. However, when variables are systematically related, it is important to consider the *possibility* that one may cause the other. Alternatively, the two correlated variables both may be the effects of some third, unidentified variable.

An example: In the United States over most of the twentieth century, the more industrial states have tended to offer stronger support for liberal Democratic candidates. This general pattern connects an independent variable, percentage of the state's adult population employed in industry, to a dependent variable, percentage of a state's electorate voting for liberal Democratic candidates. A causal relationship can be inferred from the correlation between these two variables: Conditions associated with having a lot of industry (such as urbanization, unionization, and so on) generate a preference for the liberal candidates among the people affected by these conditions. The explanation of liberal voting based on this evidence thus may emphasize the impact of industrial conditions on people's interests and the translation of these interests to a preference for liberal candidates. The causal images behind correlations are central to the representations of social life that quantitative researchers construct.

Generally, quantitative social researchers identify causation with explanation. Once the causes of a phenomenon have been identified, it has been explained. The usual sequence is:

1. a pattern of covariation is identified and the strength of the correlation is assessed,
2. causation may be inferred from the correlation, and, if so,
3. an explanation is built up from the inferred causal relationship.

Another way of understanding this is simply to say that quantitative social researchers construct images by examining patterns of covariation among variables and inferring causation from these broad patterns.

Testing Theories

While quantitative researchers often construct explanations and images from the broad patterns that they observe (like the rough correlation between income levels and educational levels) and relate these evidence-based images to their ideas about social life, they also test ideas drawn directly from social theories. Recall from Part I of this book that all social researchers are involved in long-standing, abstract conversations about social life. Social researchers use this body of thought whenever they construct images, but they also seek to advance this body of thought and to construct formal tests of ideas drawn from it.

Testing an idea is different from *using* an idea to help make sense of some pattern in a set of data or body of evidence that already has been collected. When an idea is tested, it is first used to construct an image that is based on the ideas themselves, not the evidence. The researcher constructs a theoretical image. Researchers use these theoretically based images to derive testable propositions (also called hypotheses) about evidence that has not yet been examined. Once examined, the evidence either supports or refutes the proposition (see Chapter 1).

This formal assessment of hypotheses helps social scientists determine which ideas are most useful for understanding social life. An idea that consistently fails to win support in these formal tests will eventually be dropped from the pool of ideas that social scientists use. Ideas that consistently receive support are retained.

One theoretical image in the study of social inequality is the idea that advanced societies are *achievement* oriented—they reward performance, while less advanced societies are *ascription* oriented—they reward people for who they are (for example, their family's social status). Thus, in an achievement-oriented society, a person of great ability from a low-status, impoverished background should nevertheless be successful. By contrast, in an ascription-oriented society, people born into high-status families will be successful, regardless of their talents.

These are theoretical images. There is no society that is totally achievement oriented, nor is there any society that is totally ascription oriented. However, these theoretical images have implications for inequality in the United States, which is generally considered to be an advanced society (despite its absurdly high homicide rate). Has the United States become more achievement oriented over the last forty years? Is it easier today for a talented person from a low-status, impoverished background to succeed than it was in the 1950s? The theoretical images just described link the ascendance of the achievement orientation to societal advancement, suggesting that over the last forty years it should have become easier in the United States for a talented person from a low-status background to get ahead.

Thus, the testable proposition is that evidence on "social mobility" (the study of who gets ahead) should support the idea that achievement has become more important and ascription less important in U.S. society. The increased importance of achievement criteria might be discernible in the strength of the relationship between educational achievement and subsequent income. Is the correlation between these two variables stronger in 1994 than it was in 1954? The decreased importance of ascription might be visible in the strength of the relationship between race and income. Is being black less of a liability in 1994 than it was in 1954? Of course, it would be possible to examine the effects of a variety of achievement and ascription variables on income over the last forty years (and at various points within this span of time) because there have been many surveys conducted over this period with data relevant to the proposition.

The quantitative approach is very useful for testing theoretical ideas and images such as these. Notice that these ideas are *general*—they are relevant to many cases, and they are *parsimonious*—they concern the operation of only a few causal variables. When theoretical ideas are relevant to many cases, like ideas about ascription versus achievement, we have more confidence in a test when it includes a very large number and a wide range of cases.

Making Predictions

Another goal of social research that mandates examination of large numbers of cases is making predictions. In order to be able to make predictions it is important to have as many cases as possible and to have a variety of cases. When predictions are based on many cases, researchers have the largest possible data base at their disposal and are capable of making the most accurate predictions.

For example, to predict whether middle-aged, middle class, white, Southern males will favor the Republican candidate in the next presidential election, it is necessary to know how people with this combination of characteristics generally vote in presidential elections. Do they always favor Republican candidates? Do they vote differently when the Democratic candidate is a Southerner? When issues related to national defense are important, are they more enthusiastic in their support for the Republican candidate? Clearly, the greater the volume of evidence on the political behavior of males in this category, the more precise the prediction for a future election.

Having a lot of evidence makes it easier to forecast future behavior. Knowledge of general patterns also helps. Suppose a researcher wants to predict the political behavior of middle-aged, middle class, Southern white males in an election that pits a Democratic candidate from the South against a Republican candidate who favors greater military spending. Suppose further that this particular combination of candidate characteristics has never occurred before. How can social scientists extrapolate when one condition (Democratic candidate from the South) decreases this group's support for the Republican candidate, while the other (a pro-military posture) increases its support?

Accumulated knowledge of general patterns helps in these situations. If research shows that, in general, the personal characteristics of a candidate (for example, being a Southerner) matter more to voters than the positions a candidate takes (for example, being pro-military), then the prediction would be that the Southern factor should outweigh the military factor.

Knowledge of general patterns helps social researchers sharpen their predictions by providing important clues about how to weight factors accurately, even in the face of many unknowns and great uncertainty. Because it is well suited for the production and accumulation of knowledge about general patterns, the variable-based approach offers a solid basis for making such predictions.

Contrasts with Qualitative and Comparative Research

When social researchers construct images from evidence, they may use any number of cases. Qualitative researchers typically use a small number of cases (from one to several handfuls); comparative researchers use a moderate number; and quantitative researchers use many (sometimes *thousands*). The images that qualitative researchers construct are detailed and in depth; the images that quantitative researchers construct are based on general patterns of variation across many, many cases. These general images link variation in one attribute of cases to variation in other attributes. The patterns of covariation between two or more such variables across many cases provide the basic raw material for the images that quantitative researchers construct.

The quantitative strategy favors **generality.** A quantitative researcher might show that there is a link between variation in income levels and variation in educational levels in a large sample of U.S. adults. This pattern of covariation evokes a general image of how people in the United States get ahead. If income levels covary more closely with educational levels than they do with other individual-level attributes (such as age, race, marital status, and so on), then it appears that success in the educational system is the key to subsequent material well-being. This image of how income differences arise in U.S. society is very different from one that links differences in income levels to differences in other attributes such as skin color. A key question in the application of the quantitative approach is the strength of the correlation of different causal variables, like educational level and skin color, to dependent variables, like income.

The quantitative approach prizes not only generality, but also **parsimony**—using as few variables as possible to explain as much as possible. In a study of income levels, for example, the main concern of the quantitative researcher would be to identify the individual-level attributes with the strongest correlation with income levels. Is it educational levels? Is it age? Is it parents' income? Is it skin color? Which variables have the strongest links with differences in income? By identifying the variables with the strongest correlations, quantitative researchers pinpoint key causal factors and use these to construct parsimonious images.

Parsimony and generality go together in quantitative research. Images that are general also tend to be parsimonious. It is clear that parsimony is not a key concern of the qualitative approach. Qualitative researchers believe that in order to represent subjects properly, they must be studied in depth—to uncover nuances and subtleties. Comparative researchers lie halfway in between on the issues of parsimony and generality. Rather than focus on patterns that are general across as many cases as possible—the primary concern of the quantitative approach, comparative researchers focus on diversity, on configurations of similarities and differences within a specific set of cases.

This difference between quantitative and comparative research is subtle but important. A parsimonious image that links attributes across many cases assumes that all cases are more or less the same in how they came to be the way they are. The person with low education and low income is, in this view, the reverse image of the person with high education and high income. They are two sides of a single coin.

The comparative approach, by contrast, focuses on diversity—how different causes combine in complex and sometimes contradictory ways to produce different outcomes. Thus, instead of focusing on attributes that covary with differences in income levels, like educational levels, the comparative researcher might focus on the diverse ways people achieve material success, with and without education, and contrast these with the diverse ways they fail to achieve success. From a comparative perspective, it is not a question of which attributes covary most closely with income levels, but of the different paths to achieving material success.

Of course, the comparative approach is best suited for the study of a moderate number of cases, not for the study of income differences across thousands of cases. Like the qualitative approach, the comparative approach values knowledge of individual cases. The important point in this contrast between the quantitative approach and the comparative approach is the difference between looking for variables that seem to be systematically linked to each other across many cases (a central concern of the quantitative approach) and examining patterns of diversity (a major objective of the comparative approach).

The Process of Quantitative Research

The quantitative approach is the most structured of the three research strategies examined in this book. Its structured nature follows in part from the fact that it is well suited for testing theories. Whenever researchers test theories, they must exercise a great deal of caution in how they conduct their tests so that they do not rig their results in advance. Human beings are reactive creatures. There is a large body of research showing that when people are interviewed, their responses are shaped in part by the personal characteristics of the interviewer (such as whether the interviewer is male or female). If they know what a social scientist is trying to prove, they may try to undermine the study, or they may become overcompliant. Tests in any scientific field that are not conducted carefully cannot be trusted.

The more structured nature of quantitative research also follows from its emphasis on variables. Variables are the building blocks of the images that quantitative researchers construct. But before researchers have variables that they can connect through correlations, they must be able to

specify their cases as members of a meaningful set, and they must be able to specify the aspects of their cases that are relevant to examine as variables. In short, much about the research tends to be fixed at the outset of the quantitative investigation.

This orientation contrasts sharply with those of the other two strategies. In qualitative research, investigators often do not decide what their case is a "case of" until they write up their results for publication (see Chapter 4). In the comparative approach, researchers assume that their cases are very diverse in how they came to be the way they are, and investigators often conclude their research by differentiating distinct types of cases (see Chapter 5). Of course, quantitative researchers are quite capable of differentiating types of cases, but their primary focus is on relating variables across all the cases they have data on.

Cases and variables can be fixed at the outset of a study—as they tend to be in quantitative research—only if the study is well grounded in an analytic frame. Thus, analytic frames play a very important part in quantitative research.

Analytic Frames in Quantitative Research

Researchers use analytic frames to articulate theoretical ideas about social life (see Chapter 3). Frames specify the cases relevant to a theory and delineate their major features. The importance of frames to quantitative research can be seen most clearly in research that seeks to test theories. Once a theory has been translated into an analytic frame, specific propositions (or testable hypotheses) about how variables are thought to be related to each other can be stated. Researchers can then develop measures of the relevant variables, collect data, and use correlational techniques to assess the links among relevant variables. Relationships among variables either refute or support theoretically based images.

A theory of job satisfaction may emphasize the match between a person's skills and talents, on the one hand, and the nature of the tasks he or she is required to perform, on the other. The basic theoretical idea is that people are happiest in their work when their job requires them to do things they are good at. Work that does not suit an employee makes the employee feel frustrated and dissatisfied, even useless. These theoretical ideas can be expressed in a frame that details employee and job characteristics relevant to job satisfaction.

To test the idea that job satisfaction is greatest when skills and duties are well matched, it would be necessary to elaborate this frame in advance of data collection. Of course, researchers should not remain ignorant of their research subjects before testing a theory. They should learn all that they can. The point is simply that the data used to test a theory is not the same as the evidence the researcher uses in developing or refining the hypothesis to be tested. To do this would be to rig the results of the test in a way that would confirm the researcher's ideas.

The frame becomes more or less fixed once theory testing is initiated. The job satisfaction frame is fixed on employees as cases, job satisfaction as the dependent variable, and the match between employee and job characteristics as independent variables. When a frame is fixed, the images that can be constructed from evidence are constrained. When the goal is to test theory, the images that can be constructed are further constrained by the hypothesis. In the job satisfaction example, if the researcher finds that the employees who are well matched in terms of skills and duties are not the ones with the highest levels of job satisfaction, then the image constructed from the evidence rejects the theoretically based frame.

Even when quantitative researchers are not testing theories, the images that they can construct from evidence are still constrained by their frames. In order to examine relationships among variables, it is necessary first to define relevant cases and variables. The examination of relationships among variables usually cannot begin until after all the evidence has been collected. Furthermore, the evidence that is collected must be in a form appropriate for quantitative analysis. There must be many cases, all more or less comparable to each other, and they must have data on all, or at least most, of the relevant variables. Thus, quantitative research implements frames directly, as guides to data collection, telling researchers which variables to measure.

From Analytic Frame to Data Matrix

In quantitative research the collection of evidence is seen as a process of filling in the data table (or **data matrix**) defined by the analytic frame. (An example of a small data matrix is presented in Table 6.1.) In the study of job satisfaction, the data on a single employee would fill one row of the data matrix, and there would be as many rows as employees. The columns of the data matrix would be the different employee and job characteristics relevant to the analysis. Thus, in quantitative research the data matrix mirrors the analytic frame.

The researcher would not fill in this matrix with data on just anyone. In a study of job satisfaction, for example, the researcher would probably want to collect data on all the employees of a particular factory or firm. (Of course, if the firm or factory were very large, the researcher would

probably collect a systematic, random sample of its employees.) In order to construct a good test of the theory, the researcher would choose a work setting with many different kinds of jobs and with employees possessing many different kinds of skills. This combination would provide a good setting for testing the idea that matching skills with duties is important for job satisfaction. If the researcher chose a work setting where everyone did more or less the same thing and had more or less the same skills, then it would not be an appropriate setting for testing the idea that matching skills with duties matters.

Thus, quantitative researchers exercise considerable care when selecting the cases to be used for testing a particular theory. The cases must be relevant to the theory, and they must vary in ways that allow the theory to be tested. When a theory is relevant to very large numbers (for example, all adults in the United States), the quantitative researcher uses a random sample of such cases (for example, every 10,000th person listed in the census). When it is not possible to use a national sample, the researcher may sample the people in a single city or region that is representative of the population as a whole.

Of course, not all social theories are about variation among individuals. Sometimes they are about other basic units—firms, families, factories, organizations, gangs, neighborhoods, cities, households, bureaucracies, even whole countries. In most quantitative research, cases are common, generic units like these. This preference for generic units follows from its emphasis on constructing broad, parsimonious images that reflect general patterns.

Measuring Variables

Quantitative researchers also exercise great care in developing measures of their variables. In the study of job satisfaction, the measurement of the dependent variable is critically important to the study as a whole. How should it be measured? Is it enough simply to ask employees to rate their degree of satisfaction with their jobs? Can employees be trusted to give honest and accurate assessments or will they worry that management is looking over their shoulders? Should the researcher also examine personnel files? Is this legal? Is it ethical? What about records on absenteeism? Is absenteeism a good measure of job dissatisfaction? What about asking supervisors to give their ratings of the people who work under them?

Not surprisingly, there is an immense literature on the problems of measuring job satisfaction, and comparably large literatures exist on the measurement of most of the many variables that interest social scientists.

Even variables that seem straightforward are difficult to measure with precision, and controversies abound. What does years of education measure? Knowledge? Job-relevant skills? Time spent in classrooms?

For example, it is clear that nations differ in wealth. Gross National Product in U.S. dollars per capita (GNP per capita) is a conventional measure of national wealth. However, GNP per capita has important liabilities. Some are technical. In order to get all countries on the same yardstick, their currencies must be converted to U.S. dollars. But the relevant exchange rates for making these conversions fluctuate daily. Thus, the rankings of countries on GNP per capita fluctuate daily. But wealth differences between countries are thought to be relatively long standing; differences induced by short-term exchange rate fluctuations are artificial.

A more serious problem: Some countries have a high GNP per capita but do not seem wealthy because most of their citizens do not live well. In the mid-1970s, for example, the GNP per capita of many oil-exporting countries skyrocketed, but living conditions in these countries were not as good as those of some poorer, non–oil-exporting countries. Thus, it is possible, at least in the short run of a decade or so, to have a high GNP per capita and relatively poor living conditions, which contradicts the idea of GNP per capita as a measure of national wealth.

A still more serious problem: Some countries have great income inequality, with a substantial class of very rich people, many poor people, and few in between. These countries may appear to be much better off than they are because on the average—which is what GNP per capita captures—conditions seem OK. But the reality may be one of widespread suffering in the face of extreme riches.

The issue of using appropriate measures is known as the problem of validity (see also Chapter 1). Do data collection and measurement procedures work the way social researchers claim? One way to assess validity is to check the correlations among alternative measures that, according to the ideas that motivate the study, should covary. For example, a researcher may believe that years of education is a valid measure of general knowledge and could assess this by administering a test of general knowledge to a large group of people representative of the population to be surveyed. If their scores on this test correlate strongly with their years of education, then the researcher would be justified in treating years of education in the survey of the larger population as a measure of general knowledge.

Researchers are also concerned about the reliability of their measures. **Reliability** generally concerns how much randomness there is in a particular measure (quantitative researchers refer to this as *random error*). For example, day-to-day exchange rate fluctuations produce randomness in the GNP per capita in U.S. dollars. The calculation of GNP per capita in U.S. dollars changes every time exchange rates change. Thus, GNP per capita calculated one day will not correlate perfectly with GNP per capita calculated the next, even though the estimates of the goods and services produced by each country are unchanged.

Consider an example closer to home: When employees are asked how satisfied they are with their jobs, their answers may reflect what happened that day or over the last few days. Ask them again in a month, and their answers may reflect what's happening then. Thus, when the measurements of job satisfaction taken one month apart are correlated, asking the same people the same question, the relationship may be weak because of the randomness induced by different surrounding events.

Researchers have developed a variety of ways to counteract unreliability. In research on job satisfaction, they might ask many questions that get at many different aspects of job satisfaction and use these together to develop a broad measure (for example, by adding the responses to form a total score for each person). More than likely, employees' responses to many of the questions will not change over one month. Thus, by adding together the responses to many related questions on job satisfaction, the researcher might develop a measure that is more reliable.

Measurement is one of the most difficult and most important tasks facing the quantitative researcher because so much depends on accurate measurement. If a correlation is weak, say between job satisfaction and a measure of the match between employees' skills and duties, is it because the theory is wrong or because the measures are bad? Is the measure of job satisfaction accurate? Is the measure of skills adequate? Is the measure of the match of employees' skills and duties properly conceived and executed? In the quantitative approach, there is no way to know for sure why a correlation that is expected to be strong comes out weak. Because researchers usually hold fast to their theories, they often blame their measures and complain about the difficulty of measuring social phenomena with precision.

Examining Correlations and Testing Theories

The examination of correlations among variables is the core of the quantitative approach, but quantitative researchers must travel a great distance before they can compute a single correlation. They must translate their theoretical ideas into analytic frames. They must choose appropriate cases. If there are many, many such cases, they must devise a sampling strategy. They must develop valid, reliable measures of all their variables. If the goal of the investigation is to test theory, they must also articulate the proposition to be tested and take great care in measuring the variables central to the proposition. And they must fill in the data matrix defined by their analytic frames, the cases they have selected, and the measures they have devised.

After all this preparation, the computation of correlations may seem anticlimactic. In qualitative research, the investigator engages ideas in every stage of the research, refining and clarifying categories and concepts as new evidence is gathered (see Chapter 4). In comparative research, a similar process of linking ideas and evidence occurs in the construction of truth tables (see Chapter 5). In quantitative research investigators must know a lot in advance of data collection. They must learn as much as they can about the theories they want to test, about their cases, and about how to measure their variables before they collect the data that will be used to test their theories. Thus, the examination of relationships among variables (the technique quantitative researchers use to construct evidence-based images) is near the end of a very long journey.

When quantitative researchers test theories, the key question is whether or not the correlations follow patterns consistent with the ideas that motivated the study. Sometimes this assessment involves the correlation between a single independent variable and a single dependent variable. In the study of job satisfaction: How strong is the correlation between job satisfaction and the degree to which employees' skills and duties are matched? Sometimes testing a theory involves comparing the strength of a correlation in different times or settings: Is educational level more strongly linked to income level in 1994 than it was in 1954? Sometimes testing involves comparing the correlations of several independent variables with one or more dependent variables: Is the effect of race on income stronger or weaker than the effect of education on income? Did the pattern change between 1954 and 1994?

What do researchers do when correlations do not support their theories? Sometimes, they simply report that the evidence does not support their theory. In other words, they report that they attempted to construct an evidence-based image consistent with some theory, but were unable to do so, suggesting that the theory is wrong. In general, however, the audiences for social science expect social life to be represented in some way in a research report. They do not expect a report of a failed attempt to construct a representation. Such reports should be more common than they are because the logic of theory testing (that is, the effort to figure out which ideas are best supported by evidence) indicates that negative findings (that is, failed representations) are very important.

More often, if the initial test of a hypothesis fails, researchers examine their evidence closely to see if there is support for their theory under specific conditions. After finding a weak correlation between job satisfaction and the degree to which employees' skills and duties are matched, a researcher might consider the possibility that other factors need to be considered. Perhaps employees who have been with the firm the longest are more satisfied, regardless of how well their skills are matched to their duties. This factor would need to be taken into account when examining the relation between job satisfaction and the match of skills and duties. Generally, researchers try to use their general knowledge of their cases and their theoretical understanding to anticipate refinements like these *before* they collect their data. They may also specify additional hypotheses in advance as a way to anticipate such failures.

Using Quantitative Methods

An Introduction to Quantitative Methods

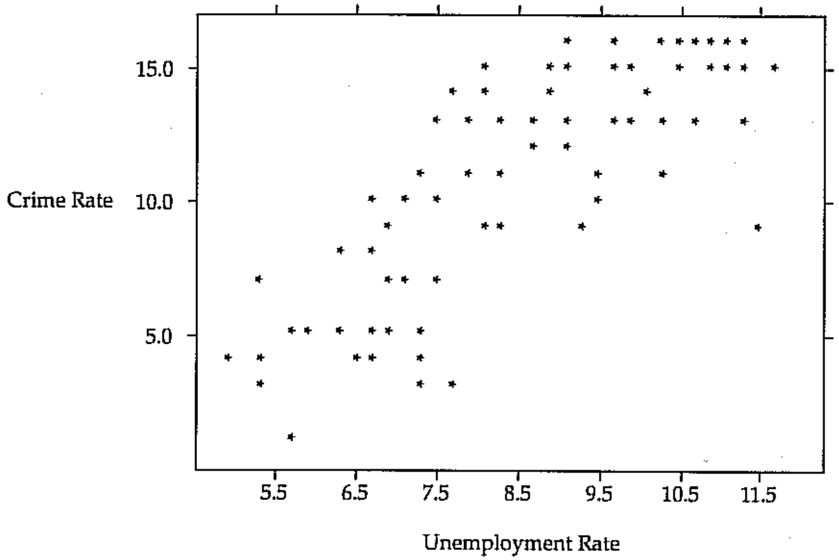
Quantitative methods focus directly on relationships among variables, especially the effects of causal or *independent* variables on outcome or *dependent* variables. Another way to think about the quantitative approach is to see the level of the dependent variable (for example, variation across countries in life expectancy) as something that *depends on* the level of other variables (for example, variation across countries in nutrition). The strength of the correlation between the independent and the dependent variable provides evidence in favor of or against the idea that two variables are causally connected or linked in some other way.

The exact degree to which two variables correlate can be determined by computing a **correlation coefficient**. The most common correlation coefficient is known as Pearson's r and is the main focus of this discussion. If the correlation is substantial and the implied cause–effect sequence makes sense, then the cause (the independent variable) is said to "explain variation" in the effect (the dependent variable).

If cities in the United States with lower unemployment rates also tend to have lower crime rates, then these two features of cities, unemployment rates and crime rates, go together; they correlate. Generally, social scientists would argue that the unemployment rate (the independent variable) explains variation across cities in the crime rate (the dependent variable). The general pattern of covariation in this hypothetical example is high unemployment rates–high crime rates, moderate unemployment rates–moderate crime rates, and low unemployment rates–low crime

FIGURE 6.1

Plot of Crime Rate with Rate of Unemployment Showing Positive Correlation



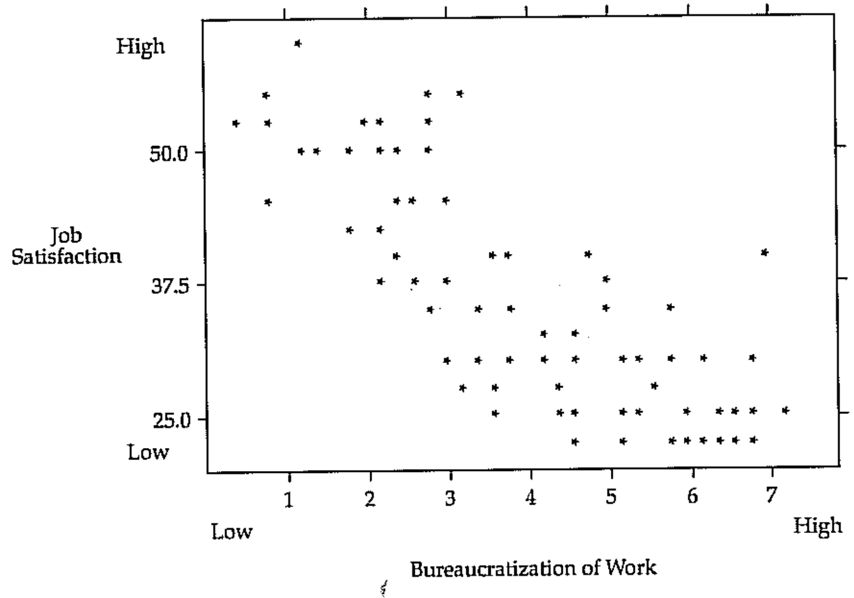
rates, as depicted with hypothetical data on cities in Figure 6.1. In this figure, the correlation is described as a **positive correlation** because high unemployment rates go with high crime rates and low unemployment rates go with low crime rates.

Some general patterns of covariation display **negative correlations.** If people who work in *less* bureaucratic settings display, on the average, *more* job satisfaction than people who work in more bureaucratic settings, then these two things, job satisfaction and degree of bureaucratization of work, are negatively correlated. This pattern can be depicted in a plot of employee data, as in Figure 6.2 which presents hypothetical evidence conforming to the stated pattern. According to the diagram, bureaucratization explains variation in job satisfaction because job satisfaction is high when people work in settings that are less bureaucratized, and vice versa.

In both examples, features of cases, called variables, are observed not in the context of individual cases, but *across* many cases. It is the pattern across many cases that defines the relation between the two features, not how the two features fit together or relate in individual cases. In the example of the positive correlation just described, it may be that one of the

FIGURE 6.2

Plot of Job Satisfaction and Bureaucratization of Work Showing Negative Correlation



cities combining high unemployment and high crime rates had a recent, dramatic increase in unemployment coupled with a decrease in its crime rate—the opposite of the general pattern across cities. (If this city's crime declined from a very high level to a merely high level, it would still appear in the high unemployment–high crime rate portion of Figure 6.1.) What happened in one case over time cannot be addressed in the correlation across many cities at a single point in time. What matters is the general pattern: Do the cities with the highest unemployment rates also have the highest crime rates? In other words, the analysis of the relation between unemployment and crime in this example proceeds across cities, not within individual cities over time.

The correlation coefficient provides a way to make a direct, quantitative evaluation of the degree to which phenomena (for example, unemployment rates and crime rates) covary across cases (such as cities in the United States). The Pearson correlation coefficient itself varies between –1.00 and +1.00. A value of –1.00 indicates a perfect negative correlation; a value of +1.00 indicates a perfect positive correlation; and a value of 0 indicates no correlation. Sometimes a finding of no correlation is important because social researchers may have strong reasons to believe that a correlation should exist. The finding of no correlation may challenge widely accepted ideas.

It is sometimes difficult to specify what value constitutes a "strong" correlation. People tend to be relatively unpredictable. Thus, some researchers consider an individual-level correlation strong if it is greater than .3 (or more negative than –.3). For whole countries, by contrast, a correlation of .3 is considered weak because many features of countries tend to be highly correlated (for example, average wealth, life expectancy, literacy, level of industrialization, rate of car ownership, and so on). When assessing the strength of correlations, it is important to consider the nature of the data used in the computation.

Computing Correlation Coefficients

The hand calculation of a correlation coefficient is time consuming but straightforward. Usually, computers are used to compute correlation coefficients such as Pearson's r. The calculation of Pearson's r is illustrated in the appendix to this book in order to show the underlying logic of the coefficient.

Remember, the goal of the computation is to assess the degree to which the values (or scores) of two variables covary across many cases, in either a positive or a negative direction. In other words, do the cases with high values on the independent variable tend to have high values on the dependent variable? Do the cases with low values on the independent variable tend to have low values on the dependent variable? If so, then a strong positive correlation exists. If high values on the independent variable tend to be associated with low values on the dependent variable, and vice versa, then a strong negative correlation exists. If there is no pattern of covariation between two variables, then there is no correlation between them.

The key to calculating a correlation coefficient is to convert the scores on two variables to Z scores, as explained in the appendix. Z scores standardize variables so that they all have the same mean or average value (0) and the same degree of variation. Table 6.1 reports data on two variables for forty countries: the average number of calories consumed per person each day (the independent variable) and life expectancy (the dependent variable). These two variables can be used to test the simple idea that in countries where nutrition is better (as reflected in more calories consumed per person) people tend to live longer (as indicated in a

TABLE 6.1

Calculating the Correlation between Calorie Consumption and Life Expectancy

Country	*Life Expectancy*	*Calorie Consumption*	*Life Expectancy Z Scores*	*Calorie Consumption Z Scores*
Niger	45	2432	–2.04	–.70
Ethiopia	47	1749	–1.85	–1.92
Mali	47	2074	–1.85	–1.34
Uganda	48	2344	–1.75	–.86
Senegal	48	2350	–1.75	–.85
Sudan	50	2208	–1.55	–1.10
Ghana	54	1759	–1.17	–1.90
Kenya	58	2060	–.78	–1.37
Zimbabwe	58	2132	–.78	–1.24
Botswana	59	2201	–.68	–1.11
Indonesia	60	2579	–.58	–.44
Morocco	61	2915	–.49	.16
Peru	61	2246	–.49	–1.03
Philippines	63	2372	–.29	–.81
Thailand	64	2331	–.19	–.88
Turkey	64	3229	–.19	.72
Syria	65	3260	–.10	.77
Brazil	65	2656	–.10	–.30
Colombia	66	2543	.00	–.50
Paraguay	67	2853	.10	.05
Mexico	69	3132	.29	.55
S. Korea	69	2907	.29	.15
Malaysia	70	2730	.39	–.17
Hungary	70	3569	.39	1.33
Poland	71	3336	.49	.91
Chile	72	2579	.58	–.44
Jamaica	74	2590	.78	–.42
Ireland	74	3632	.78	1.44
United States	75	3645	.87	1.46
Greece	76	3688	.97	1.54
Australia	76	3326	.97	.89
Spain	77	3359	1.07	.95
Italy	77	3523	1.07	1.24
Netherlands	77	3326	1.07	.89
France	77	3336	1.07	.91
Canada	77	3462	1.07	1.14
Sweden	77	3064	1.07	.43
Norway	77	3223	1.07	.71
Switzerland	77	3437	1.07	1.09
Japan	78	2864	1.17	.07

longer life expectancy). Table 6.1 also reports the Z scores for these two variables for all forty cases.

Notice that countries with high scores on life expectancy have positive scores on life expectancy Z scores, and countries with low scores on life expectancy have negative scores on life expectancy Z scores. The same is true for calorie consumption. When the Z scores for two variables are multiplied, the products indicate a lot about the correlation. If high scores on one variable correspond to high scores on the other, and low scores on one correspond to low scores on the other, then the products of the Z scores will usually be positive, indicating a positive correlation. However, if low scores on one variable generally correspond to high scores on the other, and vice versa, then the products of the Z scores generally will be negative, indicating a negative correlation.

As the appendix illustrates, when the products of pairs of Z scores for two variables are averaged over all the cases, the number that results is Pearson's correlation coefficient, a number which varies between –1.00 (perfect negative correlation) and +1.00 (perfect positive correlation). The correlation between life expectancy and calorie consumption for the forty countries in Table 6.1 is .802, a strong positive correlation. The strong covariation between these two variables is clear from simply examining the table because the countries are sorted according to their values on life expectancy. The calculation of the correlation coefficient provides a direct, quantitative assessment of the degree to which the two measures covary.

Using Correlation Coefficients

The most basic use of correlation coefficients is to assess the strength of the relation between two variables. The correlation between calorie consumption and life expectancy is strong (r = .802), suggesting that an important key to longer life expectancy is nutrition. But there are many other uses of correlations. Most of these involve the comparison of competing causes, as indicated in the strength of correlations.

Consider the correlations reported in Table 6.2. The table shows all the correlations among four variables: three independent variables (calorie consumption, GNP per capita, and doctors per capita) and one dependent variable (life expectancy). (Notice that a variable correlates perfectly with itself, as shown by the values of 1.000 in Table 6.2.) GNP per capita is a rough measure of the wealth of a country. Doctors per capita is a rough measure of the availability of medical care.

TABLE 6.2

A Correlation Matrix with Three Independent Variables and a Dependent Variable

	Dependent Variable	*Independent Variables*		
	Life Expectancy	*Calorie Consumption*	*GNP per Capita (US$)*	*Doctors per Capita*
Life expectancy	1.000	.802	.651	.721
Calorie consumption	.802	1.000	.848	.321
GNP per capita (US$)	.651	.848	1.000	.671
Doctors per capita	.721	.321	.671	1.000

The first column shows the correlations of the three independent variables with the dependent variable. Calorie consumption is the most strongly correlated with life expectancy (r = .802), followed by doctors per capita (r = .721), followed by GNP per capita (r = .651). Is it possible to conclude from this evidence that all that really matters for life expectancy is calorie consumption? In other words, if the goal is to understand the variation in life expectancy across countries, is knowing nutrition levels enough? Is it reasonable to ignore the correlations with GNP per capita and doctors per capita?

In order to answer a question like this, it is not enough simply to identify the independent variable with the strongest correlation with the dependent variable. It is also necessary to examine the correlations among the independent variables. Consider first the correlation between calorie consumption and GNP per capita. It is strong (r = .848), suggesting that countries with the best nutrition are also the richest. Given that (1) these two independent variables are strongly correlated and (2) calorie consumption has a stronger correlation with life expectancy than does GNP per capita (r = .802 versus .651), it is reasonable to conclude that the link between calorie consumption and life expectancy is more fundamental than the link between GNP per capita and life expectancy. In short, richer countries have better nutrition, but it is good nutrition that causes greater life expectancy, not wealth per se.

What about doctors per capita? The correlation between doctors per capita and calorie consumption is positive, but not strong (r = .321). Thus, in some countries nutrition may not be good, but good health care is available, while in other countries, the opposite may be the case. In other words, doctors per capita and calorie consumption are not closely linked across countries in the same way that GNP per capita and calorie consumption are. Thus, the correlation between doctors per capita and life expectancy, the dependent variable, is relatively independent of and separate from the correlation between calorie consumption and life expectancy. Even though the correlation between doctors per capita and life expectancy (r = .721) is not as strong as the correlation between calorie consumption and life expectancy (r = .802), it is an important correlation. The pattern of correlations in Table 6.2 indicates that both doctors per capita and calorie consumption affect life expectancy.

A lot can be learned from looking at a correlation matrix like the one in Table 6.2. However, some quantitative studies examine many independent and dependent variables. Quantitative researchers use advanced statistical techniques such as multiple regression analysis to disentangle correlations among independent variables and assess their separate effects on dependent variables. They also use exploratory data analysis techniques ("EDA"; see Tukey 1977) to go beyond broad patterns of covariation to identify sets of cases that deviate from these broad patterns or to uncover very subtle patterns. Sometimes these techniques can be used to identify complex patterns of causation that are specific to subsets of cases included in a study (Leamer 1978). These advanced statistical techniques are very powerful data techniques and they further the primary goals of this approach: assessing general patterns (including their limits), making projections about the future, and evaluating broad theories.

Conclusion

Quantitative methods are best suited for addressing differences across a large number of cases. These methods focus especially on the covariation between attributes that vary by level, usually across many cases. If two features of cases vary together in a systematic way, they are said to correlate. Correlation is important because it may suggest that a causal or some other kind of important relation exists between the two features that are linked. Quantitative methods provide a direct way to implement a researcher's interest in general patterns, and quantitative researchers

believe that these patterns of covariation provide important clues about social life.

In many ways, the quantitative approach appears to be the most scientific of the three approaches presented in this book. It favors generality and parsimony. It uses generic units such as individuals, families, states, cities, and countries. It can be used to assess broad relationships across countless cases. It condenses evidence to simple coefficients, using mathematical procedures. It can be used to test broad theoretical arguments and to make projections about the future. In short, it imitates many of the features and practices of hard sciences such as physics and chemistry.

While the quantitative approach does have many of the features of a hard science, it would be a mistake to portray this approach as something radically different from the other two strategies. All social research engages theoretical ideas and analytic frames, at least indirectly. All social research involves constructing images from evidence, usually lots of it. And all social researchers construct images by connecting social phenomena.