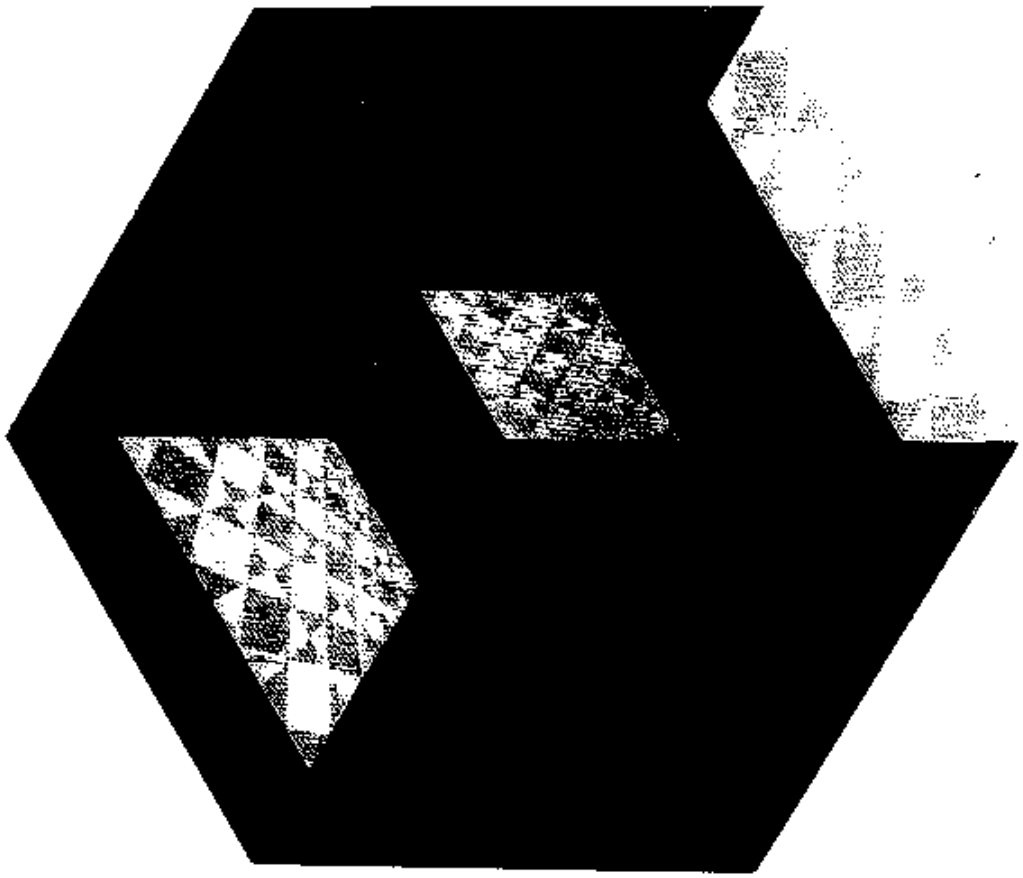
Statistics for Social Data Analysis

SECOND EDITION

George W. Bohrnstedt and David Knoke



1

The Social Research Process

The social research process often begins with questions about the world: What kind of people vote for Democrats? Do lower-income people have more children than middle-income people do? What incomes can be earned in various occupations? Why do Protestants have higher suicide rates than Catholics or Jews? Do blacks achieve less education than whites because they have lower IQs or because of other differences? Each of these questions is phrased in terms of the *relationship* between two or more observable characteristics of people or groups, such as income and occupation. We will have much to say about various relationships throughout this text, since they represent the central concept in social science.

If social research is to answer questions like these, it naturally must ask where the questions come from. Personal experience, hunch, intuition, friends' suggestions, or a variety of stimuli such as newspaper and TV accounts clearly provide relevant clues. But social scientists also have an inheritance from the past from which ideas for social research can be drawn. This inheritance is a steady accumulation of social knowledge which has been painstakingly assembled by several generations of psychologists, political scientists, sociologists, anthropologists, and economists, as well as applied researchers in education, business, and law. Together, their writings contain many theories and empirical findings about social phenomena. A student's training in the social sciences begins with an introduction to the theoretical ideas of the great masters. The

thoughts of Aristotle, Emile Durkheim, Karl Marx, Max Weber, John M. Keynes, Alfred Marshall, Charles Merriam, Arthur Bentley, Bronislaw Malinowski, Sir James Frazer, and other founding fathers of social science are a source of continuing inspiration for researchers. The importance of more contemporary if less renowned social scientists in providing ideas to be tested also must be recognized.

At its best, social science is firmly grounded in the real world. It seeks to explain social behavior, but it is distinct from related fields like social philosophy and theology which deal with idealizations that have few empirical referents. The more comprehensive social theories present distinctive views of reality which are sometimes labeled *paradigms,* or examples or patterns.[1] Another term which is used frequently is *model.* Paradigms are usually seen as broader and more encompassing than models, but both are abstractions and simplifications of the complex real world. Partitions of the seamless totality are essential if a theory is to be of any use in guiding social inquiry. No theory can seek to account meaningfully for all the significant aspects of social life. Instead, selective attention must be given to a few aspects of the phenomena to be explained. As a result, theory deals with only a part of the world and takes the rest for granted or, at least, assumes it to be sufficiently unobtrusive so it can be safely ignored while concentrating on the topic of interest.

Examples of such theoretical abstractions abound. One of the most popular in psychology is the stimulus-response, or S-R, paradigm. In B. F. Skinner's operant psychology theory, behavior is seen as purely reactive to external stimuli.[2] It posits that all behavior is a response to external stimuli, and there is no need to consider the mediating mental processes. In contrast, psychoanalytic explanations of social behavior rely on extensive, elaborate mechanisms of internal processes like Freud's trinity of id, ego, and superego. Many of these processes are unconscious, and their existence is inferred by observing patients' behaviors in dreams, slips of the tongue, and neurotic compulsions. Though the S-R and the psychoanalytic theories of behavior have markedly divergent elements, both concentrate their explanations on a few key aspects of reality and leave other features aside.

1. Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).

2. B. F. Skinner, *Science and Human Behavior* (New York: Macmillan Co., 1953).

Such analytic abstractions, we must stress, are true of all social theories. Even grandiose constructions such as the action theory of Talcott Parsons, with its elaborate exposition, necessarily focus attention on only a limited part of the near infinity of possible phenomena which could be investigated. This restriction may appear to prevent a faithful portrayal of social behavior, but it makes the research process manageable. It also is one reason science is so different from art, which attempts more complete representations of reality. If we had to take everything into account before we could say something about anything, we would never get started, and the results would be indigestible. Theoretical simplifications are a blessing which allow social scientists to state succinctly what they are going to study and what they expect to find, even though they are aware that their theses and models are incomplete representations of the real world.

1.1. What, Exactly, Is a Theory?

Up to this point, we have used the terms *theory, paradigm,* and *model* without being especially precise. You may be wondering exactly what a theory is, though no doubt you have heard casual remarks such as "I've got a theory about that," or "My theory about why they broke off their engagement is that they couldn't stand the same type of music." Such common uses of the term *theory*—as general ideas which explain events or conditions—capture only part of what social scientists mean by a theory. To the scientist, theory is a more precise explanation described in fairly abstract terms.

A **social theory** is a set of two or more propositions in which concepts referring to certain social phenomena are assumed to be causally related. While all social scientists do not accept this definition, the sentiments it expresses are fairly widely shared, and we will use it in this text.

social theory—a set of two or more propositions in which concepts referring to certain social phenomena are assumed to be causally related

proposition—a statement about the relationship between abstract concepts

concept—a precise definition of an object, behavior, perception (self or other), or phenomenon that is theoretically relevant

relationship—a connection between two concepts or variables, of either a covariational or causal nature

The definition has several components. Any theory involves two or more statements, which may be written in a common language such as English or in the more elegant shorthand languages of symbolic logic or mathematics. These statements, which are called **propositions**, are comprised of two essential elements: **concepts** and their **relationship.** In this book propositions are designated by the initial *P* and a number, as in the following example:

A Politicoeconomic Mini Theory

P1: Economic instability generates disaffection with the national political regime.

P2: Disaffection with the national regime strengthens the opposition political forces.

By our definition, these two propositions qualify as a theory, in this case a theory explaining national political cycles in terms of economic factors. Notice that the theory says nothing about the effects of other factors, such as religious revivalism or demographic changes. Additional propositions could be added to this simple theory to take these factors into account if we wished to be more comprehensive about the alleged causes of the political cycle. Propositions 1 and 2 together constitute a *minimal* theory, however. All other factors are assumed to be inoperative, *ceteris paribus* ("everything else being equal"—a phrase social scientists seem to scatter about like birdseed).

The politicoeconomic mini theory contains three concepts: economic instability, regime dissatisfaction, and opposition strength. These concepts are what the theory is all about. They are objects or activities capable of existing in more than one state or condition (that is, they are not constant). An economy can exhibit varying degrees of stability or instability, and people's dissatisfaction or opposition can cause various changes in the levels or rates. Presumably the theorist's definitions of these concepts are known to those who are reading or researching the theory. Often theoretical concepts have been extensively discussed and argued by specialists in some area, so there is some shared understanding of the terms being used. But social science concepts often have meanings that are not the same as those used in everyday language. For example, status, power, culture, and response mean quite different things to sociologists, political scientists, anthropologists, and psychologists than they do to the person in the street. To be useful in theoretical discussion, therefore, a concept must be carefully and rigorously defined. Otherwise two scientists who presume they are studying the same social phenomenon may in fact be looking at two quite different things. Obviously such a situation does not promote an orderly accumulation of facts about the social world.

But a theory is more than just a list of concepts. Our definition says that a theory describes the relationship between component

concepts, and our two propositions depict the relationships in an implicit, *causal* fashion. First, economic turmoil increases the level of dissatisfaction with national officeholders. As alienation grows, the political groups in opposition gain more support (presumably from the disaffected populace). Note that P1 and P2 have one concept in common—dissatisfaction—which links the propositions together and allows us to make a simple **deduction**. This is:

deduction—process of deriving a conclusion about relationships among concepts by logical reasoning about their connections to common concepts

P3: Economic instability increases the strength of the political opposition.

This ability to trace the connections between concepts makes theoretical propositions more than a haphazard assemblage of sentences. While every concept cannot be linked deductively with all others, every proposition within a theory should be connected to at least one other, so they form an interconnected complex.

A theory is often intended to apply only to certain conditions or objects. A theory's **scope conditions** include the time and place within which the phenomena supposedly occur. Marx's class struggle explanation is most relevant to industrial capitalism, for example. Cultural diffusion models may be limited to Southwestern pre-Columbian societies, but the microeconomic theory of the business corporation can be adapted to nonprofit organizations. A theory's **units of analysis** may be societies, nations, complex organizations, communities, statistical aggregates, small groups, families, individuals, or even personality traits. Propositions intended to hold at one level of analysis may not apply at a different level. For example, among individuals high education levels generally increase the rate of voting in elections. But in comparing communities, inner-city precincts, where the people are generally less well educated, may have higher voting rates than the suburbs, where the people are better educated, perhaps because the city political machine organizes registration and voter turnout. Theories are unfairly tested if research is conducted at the incorrect level of analysis or in an unintended time or place. All may be fair in love and war but not in social research.

scope conditions—the times, places, or units of analysis for which a social theory's propositions are expected to be valid

unit of analysis—the general level of social phenomena that is the object of observation (e.g., individual, nation)

Some social science theories are quite formal systems consisting of axioms, theorems, and logical deductive propositions.[3] Such systems are relatively rare in the social sciences, however, compared to their proliferation in the natural sciences, where the

3. Peter M. Blau, *Inequality and Heterogeneity: A Primitive Theory of Social Structure* (New York: Free Press, 1977).

core of theory often consists of extensive sets of logical, mathematical equations. Frequently social science theory is embedded in discursive essays written in everyday language. Superficially, these presentations appear to be directly accessible to ordinary readers of the English language, but on closer inspection such theories often are meaningless jargon to anyone who has not become acquainted with the background assumptions shared by social science professionals. Despite the jargon, a social theory must be capable of being put into proposition form if it is to qualify under our definition of an acceptable social theory. Any writings that do not attempt to order social science concepts about some phenomena into a coherent set of relationships cannot be considered a theory. Whether a theory is relevant to research depends on further considerations, however.

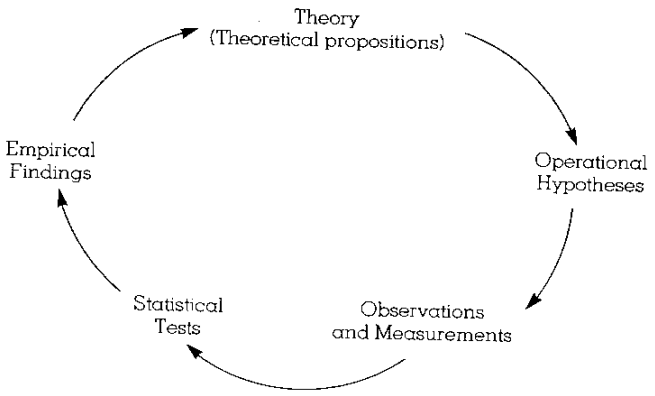
1.2. Theoretical Propositions and Operational Hypotheses

operational hypothesis —a proposition restated with observable, concrete referents or terms replacing abstract concepts

A social theory does not necessarily contain any information about how to verify its propositions. As the idealized depiction of the research cycle in Figure 1.1 shows, the next step is to generate **operational hypotheses** or working hypotheses from theory. In the social science literature the distinction between propositions and hypotheses is not always definite, and sometimes the two terms are used interchangeably. In this text the term *proposition* is used when the concepts referred to are *abstract,* and the terms *operational hypothesis, working hypothesis, research hypothesis,* or, simply, *hypothesis* are used when the concepts referred to are *concrete.* In the politicoeconomic mini theory introduced above, the three concepts are fairly abstract ideas which would permit a great deal of leeway in deciding how to detect the amount of economic instability, regime dissatisfaction, and political opposition present in a social system. Hypothesis generation restates the theoretical propositions in a form that is more relevant to actual observation in the particular social setting under study.

For example, in the United States we might decide that a good indicator of economic instability is the annual rate of inflation. Similarly, regime dissatisfaction might be indicated by the electorate's expressed attitudes about the president's economic policies, and opposition strength might mean electoral support for the party not in power in the executive branch. Of course, alternative indicators of the key concepts could be found (e.g., unemployment instead of inflation). And, in a different social system such as an

FIGURE 1.1
The Research Cycle



Source: Adapted from Walter Wallace, *The Logic of Science in Sociology* (Chicago: Aldine Publishing Co., 1969).

autocratic regime, the three concepts might have different empirical referents (e.g., terrorist bombings could be used as the indicator of opposition strength). Clearly, there are many ways to turn theoretical propositions with unobservable concepts into hypotheses with empirical referents. But the process is by no means straightforward and mechanical. Indeed, much of the argumentation among social scientists flares up over the best way to formulate hypotheses for a given theory. However, *if there are several indicators of a single concept, and one is presumably as good as another, then hypotheses verified (or falsified) by one set of indicators should also be verified (or falsified) by other sets as well.* It should be clear to you why the status of a hypothesis is strengthened or weakened by using multiple indicators of its concepts.

The theoretical propositions of our theory can be restated as operational hypotheses, using the indicators designated above. Using the initial *H* instead of *P*, these hypotheses would be stated as follows:

H1: Higher rates of inflation generate negative attitudes toward the president's economic policies.

H2: Greater dissatisfaction with the president's economic policies increases the electoral support for the party not in the White House.

These hypotheses now give us something concrete to examine in the research. We are told not only what specific things to look at but how these indicators relate to each other. To be useful in promoting research on social behavior, a theory's *propositions must be translatable into testable hypotheses,* that is, statements about the relationships between *observable* phenomena. A theory whose concepts are so vague that meaningful indicators of them cannot be found would be useless for research purposes. For this reason operant psychology has generated more testable and researchable hypotheses than the psychoanalytic approach, to return to an example used above.

We would not wish to dignify some writings with the designation *theory.* Medieval scholasticism was basically a system of thought which was *atheoretical,* or without theory, and dealt with the relationships between nonobservable phenomena, such as the famous question, "How many angels can dance on the head of a pin?" While social science has become increasingly adept at spotting and weeding out such nontheories, you will often encounter such writings in your work. When you do, you should ask yourself whether the author presents a useful theory: Are the concepts related to each other, and do observable indicators exist for the more general, abstract concepts?

variable—a characteristic or attribute of persons, objects, or events that differs in value across such persons, objects, or events

constant—a value that does not change

A term which is frequently applied to the indicators in a hypothesis is *variable.* A **variable** is a characteristic of persons, objects, or events that differs in value across persons, objects, or events. Anxiety is a variable: People differ in their degrees of anxiousness. Income is a variable: People beg, borrow, steal, or earn the money they need. *Anything that does not vary cannot be a variable;* instead it is a **constant**, or a single category of a variable. Female is not a variable but a constant. But sex (gender) is a variable because there are two states into which observations can be classified.

1.3. Independent and Dependent Variables

The research scientist is interested in establishing regularities, in particular, regularities that are due to causal relationships between and among variables. These relationships are based on the assumption that change in one variable results in some predictable change in another variable. We can think of one variable as being an antecedent and another as its consequent. The antecedent

variable (or variables) is called the **independent variable**, and the consequent variable is called the **dependent variable**.

independent variable— a variable that has an antecedent, or causal, role in relation to a dependent variable

dependent variable— a variable that has a consequent, or affected, role in relation to an independent variable

In its everyday meaning, the term *independent variable* is more applicable to experiments where antecedent variables can be manipulated freely by the researcher. An agricultural experimenter, for example, can freely vary the amount of fertilizer applied to various plots (independent variables) and observe the growth consequences of plants (dependent variable) for various amounts of fertilizer applied. In contrast, most of the outcomes in the social sciences can be assumed to be fixed for a given observation. A person who is selected for a study sample has a fixed religious identification, sex, and so on. But even though many variables thought to be antecedent cannot be manipulated in the social sciences, the term *independent variable* is still employed.

The terms *independent* and *dependent* are widespread in the research literature of the social sciences, and you should be certain to learn the distinction between them. Some examples may help to reinforce the distinction. If it is assumed that people's yearly earnings vary by educational attainment, then earnings is the dependent variable (consequent) and educational attainment the independent variable (antecedent). Similarly, religiosity varies by categories of religious identification; the independent variable is religious identification and the dependent variable is religiosity. For such independent variables, a statement that "A change in X creates a change in Y" makes no sense.

Variables that cannot be manipulated are sometimes called **status variables**. A person's race simply cannot be changed, for example. Nevertheless, different positions on a status variable (e.g., race) can often be related to different outcomes on another variable (e.g., yearly earnings). Thus even though race cannot be manipulated, it still makes sense to think of it as an independent variable. A better approach to status variables, instead of thinking in terms of changes in one variable producing changes in another, is to think in terms of various outcomes of the dependent variable as a function of the various conditions of the independent variable.

status variable— a variable whose outcomes cannot be manipulated

> The difference between an independent variable and a dependent variable is specific to a given hypothesis. Thus, in our example on the relationships among political and economic variables, attitude toward the president's policies is the dependent variable in H1,

since it is being explained. But attitude is the independent variable in H2, since it explains, or is antecedent to, electoral support.

1.4. Rejecting Hypotheses

The reason for stressing the production of operational hypotheses from theoretical propositions is simple but fundamental: *Unless it is possible to reject erroneous statements about the real world, it will not be possible to advance our theoretical knowledge of social behavior.* The idea of rejecting or falsifying propositions is an important one, since, strictly speaking, *it can never directly be proven that a proposition is true.* To try to do so would imply that the results of a single research study would hold across all time, all persons, and all cultures—and this simply cannot be known from a single study. Therefore "truth" is a goal which can be approached only incrementally. A proposition that has failed to be falsified in many studies across time and in several different cultures has more truth value than one which has been supported by only a single study. When a proposition has been falsified in a study, the implication is that the proposition is not true in general, although it *might* be true for certain groups at certain times. While social scientists hope that in the social world there are propositions that hold across time and space, they work in such an imprecise science that they cannot yet be certain that universality is the case. The important point is that falsification is easier than verification in science, and for this reason truth is approached only indirectly.

As testable statements about relationships between observable variables, hypotheses provide a capability to reject propositions. If the variables in a hypothesis are acceptable referents for the concepts of theoretical propositions, the results of research on relationships between the variables will tell something about the truth value of the theoretical propositions from which the hypotheses originally came. If in empirical analysis the hypotheses are found to be incorrect, the truth of the theory will be in doubt, at least insofar as the situation investigated is concerned. If, on the other hand, the findings are consistent with the hypothetical expectations, our belief will be strengthened that the theory helps explain something about the social world.

Imagine a theory about the effect of a certain type of teaching method on pupils' learning abilities. Suppose a well-designed

experiment finds that students taught under the new method do better on a standard test of knowledge than students taught with several other methods. These results might be obtained across many replications of the experiment. Still the researchers could not conclude that the learning theory from which these hypotheses were generated has been conclusively proven. A single instance in which the new instruction method failed to yield the hypothesized outcome would cause its unqualified validity to be rejected. And, since even a very extensive experimental sequence cannot examine all possible combinations of students, subject matter, teacher experience, and so on, the possibility remains—no matter how remote and improbable—that the hypotheses one day might not be supported.

Later in this text we will show how the probabilities of error are taken into account in statistical tests. Here we are concerned with the implied strategy of hypothesis testing. Research cannot be conducted under all the relevant conditions for all the relevant populations (it is not possible to do research on the past or the future, for example), and, increasingly, the funds available for social research are being limited. Scientists would be foolish to adopt as a criterion the *proof* of a testable theory from evidence used to test hypotheses. It is better to adopt a criterion that can produce hypotheses capable of *rejection* under clearly specified conditions.

This strategy of building research foundations on hypothesis refutation was promoted by a philosopher of science, Karl Popper; you can read his arguments in the original.[4] Popper's ideal theory is one that clearly spells out in advance exactly what sort of data and results it would be necessary to find in order to reject its claims. In the natural sciences such conditions are fairly easy to find, and there is a lengthy history of theoretical progress through *critical experiments*. In these tests, the hypotheses produced by competing theories can be lined up in direct contradiction to each other, the experimental or observational data relevant to the test can be collected, and the results will refute all but one of the hypotheses. The findings can strike a devastating blow against entire theories from which the erroneous hypotheses were deduced. For example, the solar eclipse of 1911 provided evidence against Newton's theory but was consistent with Einstein's theory of the universe.

4. Karl Popper, *The Logic of Scientific Discovery* (New York: Basic Books, 1959).

Despite the appealing simplicity of this falsification approach to scientific rationalism, its application to social science has been open to question. The reasons are many, including imprecise concepts, propositions that use concepts without empirical referents, theories of uncertain scope or coverage, focuses on different levels of explanation, and the often debatable quality of measurement of the critical variables.

> Despite these problems, we hold to the idea that the cumulation of reliable, useful knowledge in the social sciences depends on hypotheses for which refutation is sought. In sum, hypotheses should be stated so that they are capable of rejection by empirical evidence.

1.5. Operationalization

After deciding on a theory's concepts, the researcher must *operationalize* them into a set of observable variables. An **operation** is a method for observing and recording those aspects of persons, objects, or events that are relevant to testing a hypothesis.

operation—any method for observing and recording those aspects of persons, objects, or events that are relevant to testing a hypothesis

Researchers have invented and developed a wide variety of operations to meet different research needs. An entire volume would be required to detail the various approaches—historical records and archival documents; pottery shards and artifacts; participant observations; survey and questionnaire responses; content analyses of verbal exchanges; unobtrusive observations; videotapes; automatic experimental tabulations; brain-wave and galvanic skin responses. The list is potentially endless, limited only by social scientific imagination and inventiveness. These operations are all routinized, widely shared sets of activities. The forms in which observations are recorded are quite diverse, ranging from the rambling narratives of raw field notes in participant-observer studies to the highly specialized numerical coding schemes of small-group interactions in laboratory settings.

Establishing linkages between concepts and their empirical referents is a critically important step in research. Researchers want to be certain that the operations have validity, that is, they result in *valid measures* of a theory's concepts. **Validity** refers to the degree to which an operation results in a measure that accurately reflects the concept it is intended to measure.

validity—the degree to which an operation results in a measure that accurately reflects the concept it is intended to measure

A synonym for validity is *accuracy*. To the degree that an operation results in observable measures that are accurate representations of a theory's concepts, the resulting measures are said to be *valid*.

In the politicoeconomic mini theory introduced in Section 1.1, the first proposition is:

P1: Economic instability generates disaffection with the national political regime.

The two concepts in this proposition are *economic instability* and *disaffection with the national regime*. The operational hypothesis later suggested as a test of P1 is:

H1: Higher rates of inflation generate negative attitudes toward the president's economic policies.

The concept *economic instability* is operationally measured in H1 by *rates of inflation* and the concept *disaffection with the national political regime* is measured by *negative attitudes toward the president's economic policies*. The concept rates of inflation is valid to the degree it accurately measures economic instability. And the concept negative attitudes toward the president's economic policies is valid to the degree it accurately measures disaffection with the national regime.

While the concept of validity may seem to be a relatively straightforward one, it is more complex than meets the eye. Consider the following simple example. Let's assume that your weight does not change at all on a given Monday through Friday. On Monday through Thursday when you get on your bathroom scale it registers 170. Much to your surprise, when you weigh yourself on Friday the scale indicates that you weigh 77.3! You scream to your roommate that the scale is broken—it is not validly measuring your weight. Then you find out that your roommate has substituted a scale that measures in kilograms for the regular scale which measures in pounds. Since a kilogram is equivalent to 2.2 pounds, the scale *is* valid since it is in fact measuring the concept it is intended to measure—weight. If the second scale had been measuring percent body fat, then it would have been invalid as a measure of weight.

The topic of validity is extremely important in research, but a discussion of the methods for assessing validity is beyond the

scope of this book. Good discussions can be found in other texts.[5]

reliability—the degree to which different operations of the same concept yield the same results

Researchers must have measures that are *reliable* as well as valid. **Reliability** refers to the degree to which different operations of the *same* concept yield the same results.

> Reliability refers to the degree to which observations of a study are *repeatable*. A measuring instrument is said to be reliable according to the degree to which it generates *consistent* observations at two points in time. Or a measure is reliable to the degree that two different researchers using the same instrument on the same sample would generate the same observations.

Another common approach is to use several similar but not identical measures of the same concept. Industrialization, for example, might be measured by using various countries' gross national products and kilowatt hours consumed. Or people's religiosity might be measured by ascertaining their belief not only in an Almighty God but also in a life after death. To the degree that countries on the one hand and people on the other are consistently rank ordered using the two different indicators, the measurement approach is said to be reliable.

> If the observations in social research are highly unreliable, the ability to make any meaningful conclusions about the validity of social theories is seriously impaired. *Indeed, a measure of a concept cannot be valid if it is unreliable. However, a measure can be reliable (or consistent) without being valid (or a true measure).*

1.6. Measurement

Observations and measurements are distinct processes, though we show them together in the diagram of the research cycle in Figure 1.1. Once a researcher has determined how to operational-

5. Recommended references are Nan Lin, *Foundations of Social Research* (New York: McGraw-Hill Book Co., 1976), and Jum C. Nunnally, *Psychometric Theory* (New York: McGraw-Hill Book Co., 1967).

ize the variables in the hypotheses under investigation and what sample of units to observe, he or she must decide how numbers are to be assigned to the information collected so that statistical interpretation of the findings is possible. The assignment of numbers to observations according to a set of rules is called **measurement**. By translating systematic observations into a numerical system, measurement provides a means whereby mathematical manipulations can be used to assess the validity of hypothesized relationships between variables. The measurement process is one more step in the reduction of complex reality to more manageable simplicity.

measurement—the assignment of numbers to observations according to a set of rules

The types of numbers which can be assigned to observations depend on the (assumed) underlying characteristics of the phenomena to be studied. There are two broad classes of measurements: discrete and continuous.

Discrete variables classify persons, objects, or events according to the kind or quality of their attributes. The simplest type of discrete measure is the *dichotomy,* in which observations are classified only according to whether the defining attribute is present or absent. Examples are: dead or not dead; male or not male; black or not black. Other discrete attributes are *multicategorical* (or polychotomous); that is, there are more than two outcomes. Examples are: hair color classified as blond, brunette, or redhead; country of national origin; county of residence. A further distinction can be made among discrete variables according to whether or not their outcomes are orderable. **Orderable discrete variables** have outcomes that can be ranked from low to high, or vice versa. An example of an orderable discrete variable is "How well is President Reagan doing his job?" The response categories might be: "excellent, good, fair, or poor." **Nonorderable discrete variables** cannot be ranked; examples include race, ethnicity, sex, and national origin.

discrete variable—a variable that classifies persons, objects, or events according to the kind or quality of their attributes

orderable discrete variable—a discrete measure in which the categories are arranged from smallest to largest, or vice versa

nonorderable discrete variable—a discrete measure in which the sequence of categories cannot be meaningfully ordered

Continuous variables classify persons, objects, or events according to the magnitude or quantity of their attributes. The principal difference between continuous and discrete variables is that *continuous measures allow for fractional numeric values, whereas discrete measures do not.* In the natural sciences continuous variables abound: weight, height, time, velocity, and so on. In the social sciences many concepts also are assumed to be continuous; industrialization, physical attractiveness, and political liberalism are three examples. While the concepts are clearly continu-

continuous variable—a variable that, in theory, can take on all possible numerical values in a given interval

ous, however, their operational measures often only approach this goal. As a result, considerably more measurement error occurs in the social sciences. Social research must rely on crude questionnaires, rating scales and the like, because the available instruments nowhere approach the accuracy of the beam scales and electron microscopes used in the physical sciences. Nevertheless, percentages and proportions do serve as continuous measures. The percentage of 18 year olds who have ever smoked marijuana, the rate of bankruptcy among small businesses, and the proportion of registered voters who actually vote in a given election are examples.

mutual exclusiveness—a property of a classification system that places each observation in one and only one category of a variable

exhaustiveness—a property of a classification system that provides sufficient categories so that *all* observations can be located in some category

level of measurement—a classification of measurement scales according to the amount of information about observations recorded. This information generates four types of scales: nominal, ordinal, interval and ratio

nominal level of measurement—scale that assigns a name or number to observations in purely arbitrary sequence

ordinal level of measurement—scale that assigns numbers to observations in sequence, from lesser to greater amounts of the measured attribute

interval level of measurement—scale that assigns numbers to observations that reflect a constant unit length between categories

ratio level of measurement—scale that assigns numbers to observations to reflect the existence of a true, or absolute, zero point

With continuous and discrete variables alike, observations must be assigned to categories that are *both* mutually exclusive and exhaustive. **Mutual exclusiveness** refers to the need to place each observation in one and only one of the variable's categories. **Exhaustiveness** means that sufficient categories must exist for all observations to be classified into some category.

For example, if in measuring religious identification the categories Protestant, Catholic, and Baptist were used, the mutual-exclusiveness criterion would be violated, since Baptists are Protestants as well. If only the Catholic and Protestant categories were used, then the exhaustiveness criterion would be violated, since there is no category for Jews. Better categories for this variable would be: "Catholic, Protestant, Jew, other, or not ascertained." In some cases it might be better still to break the Protestant category up into several additional categories.

While "other," "not determined," or "uncertain" are convenient and often necessary categories for the recording of observations, *statistical models assume that phenomena classified the same way are in fact identical or nearly identical in how they relate to other variables.* For this reason it is better to use refined, well-thought-out categories. As much as possible, the lumping together of disparate outcomes into a single "other" category should be avoided.

The distinction between discrete and continuous variables or measures will be maintained throughout our discussion of statistical methods. Many older textbooks on social statistics devoted considerable attention to the idea of four **levels of measurement: nominal, ordinal, interval, and ratio scales.** This approach to measurement was popularized decades ago by the

psychophysicist S. S. Stevens,[6] and for some time data analysts were urged to use only certain types of statistics with data measured at each of the different levels. In the approach we have adopted for this text, the levels are considered important to measurement theory, but they have no relevance to the proper choice of statistical techniques to be used to analyze social data. In modern practice, data analysts ignore the levels in favor of the simpler discrete-continuous distinction. In the brief discussion of the four measurement scales in Box 1.1, they are identified as discrete or continuous (or, in the case of ordinal scales, as both). *The distinction between discrete and continuous variables is the most meaningful criterion for the selection of statistical techniques.*

1.7. Statistical Tests

After operationalizing the measures and making observations, the researcher's next step in the research cycle (as shown in Figure 1.1) is the performance of statistical tests, or calculations. We will not say much about statistical analysis in this chapter, since the rest of the text is a detailed examination of various techniques. As a brief introduction, however, three broad categories of statistical usage can be distinguished: descriptive statistics, measures of association, and inference.

1.7.1. Descriptive Statistics

Single variables may be *described* according to their numerical properties. Newspaper and TV accounts use such figures as *averages* (mean rainfall for June; average take-home pay for workers), *rates* (annual unemployment rate; monthly rate of inflation), *proportions* (percentage of blacks earning over $30,000; proportion of first marriages ending in divorce), and *frequency counts* (populations of the largest states; number of intercontinental missiles stockpiled by major powers). Such **descriptive statistics** form the bedrock for more advanced techniques.

descriptive statistics—numbers that describe features of a set of observations; examples are percentages, modes, variances, and correlations

1.7.2. Measures of Association

The core of our approach to presenting statistics as tools for data analysis is the analysis of the *relationships* between two or more

6. S. S. Stevens, *Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects* (New York: John Wiley & Sons, 1975).

BOX 1.1 Levels of Measurement

S. S. Steven's levels of measurement (which he called *scales*) are, in sequence from lowest to highest: (1) nominal, (2) ordinal, (3) interval, and (4) ratio. In general, the higher the level of the scale, the greater the information about the differences among variables measured by that scale.

A *nominal scale* assigns names or numbers to classes of outcomes in a purely arbitrary sequence. For example, the home states of students in your class can be presented in alphabetical or geographical sequence, with no effect on the number of students found in each category. The outcomes of a nominal scale are not inherently orderable.

The categories in an *ordinal scale* are arranged in a sequence from lesser to greater amounts of the measured attribute. No assumption about a constant difference between adjacent categories is made. Many social science attitude scales are ordinal measures which ask respondents to indicate the extent of their agreement with some statement.

An *interval scale* implies a constant unit length between categories. Any numbers assigned to the categories on this scale must reflect this distance information. The National Opinion Research Center (NORC) occupational prestige scale is a good example: The difference between two occupations rated at 80 and 60 in prestige is presumably the same difference as that between occupations rated 55 and 35, that is, 20 prestige points.

The *ratio scale* is defined by the presence of a true, or absolute, zero point: 0 on a ratio scale means no quantity of the attribute is present. People's ages or annual incomes in dollars are measured at the ratio level.

In terms of basic discrete-continuous measurement, the interval and ratio scales are continuous measures, and their numerical values can be analyzed by all statistics suitable to continuous variables. Nominal scales are clearly discrete variables, since their categories are arranged in an arbitrary sequence. Ordinal scales can be treated as discrete measures, particularly when only a handful of categories are involved. However, *since most ordinal scales are assumed to reflect an underlying concept which is continuous, their measures will often be treated as if they were continuous as well.*

variables. The relationships to be analyzed are drawn from the research hypotheses generated by a social theory. An entire branch of statistics is concerned with **measures of association** or covariation among variables. These techniques make it possible to determine the empirical support for hypotheses such as "The more years of formal schooling, the more income people earn," or "The greater the test anxiety, the lower the test performance."

measures of association—statistics that show the direction and/or magnitude of a relationship between variables

1.7.3. Inference

Much social research is performed on a **sample**, or limited number, of observations rather than observations of an entire **population**, or universe, of persons, objects, or events. Typically, the researcher collects data on some portion of all units of analysis but would like to be able to say something about the larger group from which the sample was drawn. Since selecting a sample involves a certain element of chance—who gets into or gets left out of the sample is the outcome of a probability process—it cannot be simply concluded that the results of a sample analysis hold for the population as a whole. A different sample drawn from the same population might have yielded a different set of findings. An entire statistical field addresses the problem of **inference**, or making generalizations from descriptive statistics based on sample data to the population parameters. These techniques are based on the laws of probability.

sample—a subset of cases or elements selected from a population

population—a set of persons, objects, or events having at least one common attribute to which the researcher wishes to generalize on the basis of a representative sample of observations

inference—the process of making generalizations or drawing conclusions about the attributes of a population from evidence contained in a sample

Suppose a researcher wanted to know how well you could spell, or how many of the 500,000 words in the English vocabulary you know. Obviously, you could not be asked to attempt to spell every word or give an acceptable definition of each one. The population of words is too large. But it would be possible to estimate the proportion of words in your vocabulary or the extent of your spelling skills by testing you with a **random sample** of words drawn from the dictionary. *If the sample is random and sufficiently large, the proportion of correct answers on the sample will allow an inference of how you might perform if the impossible test could have been conducted on the entire word population.*

random sample—a sample whose cases or elements are selected at random from a population

Along with descriptive statistics and measures of association, this text will expose you to the variety of appropriate inferential statistics. These statistics help determine whether the findings of

the sample tested can be generalized to the population from which the sample was drawn.

1.8. Completing the Research Cycle

Statistical tests are an indication of whether or not the assertions of the research hypotheses generated by a social theory should be rejected, in line with the idea of hypothesis refutation described above. If the empirical findings do not permit the hypothesis to be rejected, there is increased confidence that the social theory is valid, at least for the population studied. If the findings cause the hypothesis to be rejected, and there is a fairly high degree of confidence that the sample results reflect the true state of affairs in the population, belief in the validity of the theory will be severely shaken. Indeed, the researcher may want to modify the theory to take into account the new findings. Such modifications, in effect, produce a new social theory, although the family resemblance to the now-discredited former theory may still be strong. This feedback of research findings into social theory completes the cycle which was opened with theoretical propositions in Figure 1.1.

Our depiction of the research cycle is a highly idealized simplification of what may be a very complex process involving many researchers working on a single social problem, with minimal coordination of their activities. We believe this idealization captures the essential spirit of the social research enterprise, but the best way for you to find out what goes on in it is to try doing some research yourself.

> There is no simple substitute for direct experience in trying to understand what any social activity is all about. We have designed this text, therefore, to provide hands-on experience in doing research through the analysis of actual data.

If you follow our examples and attempt to replicate them on topics of your choice, you will gain a greater insight into the research process than you could by merely reading about how others have done their research. Indeed, the research reports in professional journals do not convey much about the hard work, false starts, raised expectations and shattered hopes, doldrums and break-

These hypotheses now give us something concrete to examine in the research. We are told not only what specific things to look at but how these indicators relate to each other. To be useful in promoting research on social behavior, a theory's *propositions must be translatable into testable hypotheses,* that is, statements about the relationships between *observable* phenomena. A theory whose concepts are so vague that meaningful indicators of them cannot be found would be useless for research purposes. For this reason operant psychology has generated more testable and researchable hypotheses than the psychoanalytic approach, to return to an example used above.

We would not wish to dignify some writings with the designation *theory.* Medieval scholasticism was basically a system of thought which was *atheoretical,* or without theory, and dealt with the relationships between nonobservable phenomena, such as the famous question, "How many angels can dance on the head of a pin?" While social science has become increasingly adept at spotting and weeding out such nontheories, you will often encounter such writings in your work. When you do, you should ask yourself whether the author presents a useful theory: Are the concepts related to each other, and do observable indicators exist for the more general, abstract concepts?

variable—a characteristic or attribute of persons, objects, or events that differs in value across such persons, objects, or events

constant—a value that does not change

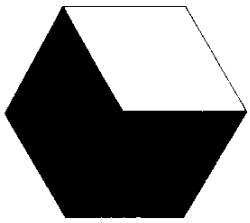
A term which is frequently applied to the indicators in a hypothesis is *variable.* A **variable** is a characteristic of persons, objects, or events that differs in value across persons, objects, or events. Anxiety is a variable: People differ in their degrees of anxiousness. Income is a variable: People beg, borrow, steal, or earn the money they need. *Anything that does not vary cannot be a variable;* instead it is a **constant**, or a single category of a variable. Female is not a variable but a constant. But sex (gender) is a variable because there are two states into which observations can be classified.

1.3. Independent and Dependent Variables

The research scientist is interested in establishing regularities, in particular, regularities that are due to causal relationships between and among variables. These relationships are based on the assumption that change in one variable results in some predictable change in another variable. We can think of one variable as being an antecedent and another as its consequent. The antecedent

throughs that are typically encountered on the way to discovering a few gems of social behavior. By concentrating on the statistical analysis of already collected data, this text exposes you to the heart of the research cycle—that point at which we find out whether the beliefs about social reality expressed in the hypotheses have any support. From our own experiences in learning to become social scientists (a learning experience that is never completed, as we found out while writing this book), *doing* social research can be an exhilarating or a frustrating experience—usually both! Though your experience in using this text may be at neither extreme, you are certain to gain considerable insight into an increasingly important aspect of our society.

Review of Key Concepts



This list of the key concepts introduced in Chapter 1 is in the order of appearance in the text. Combined with the definitions in the margins, it will help you review the material and can serve as a self-test for mastery of the concepts.

social theory
proposition
concept
relationship
deduction
scope conditions
unit of analysis
operational hypothesis
variable
constant
independent variable
dependent variable
status variable
operation
validity
reliability
measurement
discrete variables
orderable discrete variables
nonorderable discrete variables
continuous variables
mutual exclusiveness
exhaustiveness
levels of measurement
nominal scale
ordinal scale
interval scale
ratio scale
descriptive statistics
measures of association
sample
population
inference
random sample