CASE STUDY RESEARCH

Design and Methods

Second Edition

Robert K. Yin

Applied Social Research Methods Series
Volume 5

SAGE Publications
International Educational and Professional Publisher
Thousand Oaks London New Delhi

1

Introduction

The case study is but one of several ways of doing social science research. Other ways include experiments, surveys, histories, and the analysis of archival information (as in economic studies). Each strategy has peculiar advantages and disadvantages, depending upon three conditions: (a) the type of research question, (b) the control an investigator has over actual behavioral events, and (c) the focus on contemporary as opposed to historical phenomena.

In general, case studies are the preferred strategy when "how" or "why" questions are being posed, when the investigator has little control over events, and when the focus is on a contemporary phenomenon within some real-life context. Such "explanatory" case studies also can be complemented by two other types—"exploratory" and "descriptive" case studies. Regardless of the type of case study, investigators must exercise great care in designing and doing case studies to overcome the traditional criticisms of the method.

THE CASE STUDY AS A RESEARCH STRATEGY

This book is about the design and conduct of case studies *for research purposes*. As a research strategy, the case study is used in many situations, including:

- Policy, political science, and public administration research
- Community psychology and sociology
- Organizational and management studies
- City and regional planning research, such as studies of plans, neighborhoods, or public agencies
- The conduct of dissertations and theses in the social sciences—the academic disciplines as well as professional fields such as business administration, management science, and social work

This book covers the distinctive characteristics of the case study strategy, compared with other types of research. Importantly, the book deals with

design, analysis, and reporting issues—and not merely the more traditional focus on data collection or fieldwork.

The overall goal of this book is to help investigators deal with some of the more difficult questions commonly neglected by available research texts. So often, for instance, the author has been confronted by a student or colleague who has asked (a) how to define the case being studied, (b) how to determine the relevant data to be collected, or (c) what should be done with the data, once collected. This book, it is hoped, answers these questions.

However, this book does not cover all uses of case studies. For example, it is not intended to help those who might use case studies as teaching devices, popularized in the fields of law, business, medicine, or public policy (see Llewellyn, 1948; Stein, 1952; Towl, 1969; Windsor & Greanias, 1983) but now prevalent in virtually every academic field, including the natural sciences. For teaching purposes, a case study need not contain a complete or accurate rendition of actual events; rather, its purpose is to establish a framework for discussion and debate among students. The criteria for developing good cases for teaching—usually of the single- and not multiple-case variety—are far different than those for doing research (e.g., Caulley & Dowdy, 1987). Teaching case studies need not be concerned with the rigorous and fair presentation of empirical data; research case studies need to do exactly that.

Similarly, this book is not intended to cover those situations in which cases are used as a form of record keeping. Medical records, social work files, and other case records are used to facilitate some practice, such as medicine, law, or social work. Again, the criteria for developing good cases for practice are different than those for designing case studies for research.

In contrast, the rationale for this book is that case studies are increasingly used as a research tool (e.g., Hamel, 1992; Perry & Kraemer, 1986) and that you—who may be a seasoned or budding social scientist—would like to know how to design and conduct single- or multiple-case studies to investigate a research issue. This book concentrates heavily on the problem of designing and analyzing case studies and is not merely a guide to collecting case study evidence. In this sense, the book fills a void in social science methodology, which is dominated by texts on "field methods," offering few guides on how to start a case study, how to analyze the data, or even how to minimize the problems of composing the case study report. This book covers all of the phases of design, data collection, analysis, and reporting.

As a research endeavor, the case study contributes uniquely to our knowledge of individual, organizational, social, and political phenomena. Not surprisingly, the case study has been a common research strategy in psychology, sociology, political science, business, social work, and planning (Yin, 1983). Case studies are even found in economics, in which the structure of a given industry, or the economy of a city or a region, may be investigated by using a case study design. In all of these situations, the distinctive need for case studies arises out of the desire to understand complex social phenomena. In brief, the case study allows an investigation to retain the holistic and meaningful characteristics of real-life events—such as individual life cycles, organizational and managerial processes, neighborhood change, international relations, and the maturation of industries.

COMPARING CASE STUDIES WITH OTHER RESEARCH STRATEGIES

When and why would you want to do case studies on some topic? Should you consider doing an experiment instead? A survey? A history? A computer-based analysis of archival records such as student records?

These and other choices represent different research strategies. (The following discussion focuses only on five choices and does not attempt to catalog all of them, however.) Each is a different way of collecting and analyzing empirical evidence, following its own logic. And each strategy has its own advantages and disadvantages. To get the most out of using the case study strategy, you need to know these differences.

A common misconception is that the various research strategies should be arrayed hierarchically. We were once taught to believe that case studies were appropriate for the exploratory phase of an investigation, that surveys and histories were appropriate for the descriptive phase, and that experiments were the only way of doing explanatory or causal inquiries. The hierarchical view reinforced the idea that case studies were only an exploratory tool and could not be used to describe or test propositions (Platt, 1992a).

This hierarchical view, however, is incorrect. Experiments with an exploratory motive have certainly always existed. In addition, the development of causal explanations has long been a serious concern of historians, reflected by the subfield known as historiography. Finally, case studies are far from being only an exploratory strategy. Some of the best and most famous case studies have been both descriptive (for example, Whyte's *Street Corner Society*, 1943/1955; see BOX 1) and explanatory (see Allison's *Essence of Decision: **Explaining** the Cuban Missile Crisis*, 1971 [emphasis added to title]; see BOX 2).

The more appropriate view of these different strategies is a pluralistic one. Each strategy can be used for all three purposes—exploratory, descriptive, or

BOX 1
A Famous Descriptive Case Study

Street Corner Society (1943/1955), by William F. Whyte, has for decades been recommended reading in community sociology. The book is a classic example of a descriptive case study. Thus it traces the sequence of interpersonal events over time, describes a subculture that had rarely been the topic of previous study, and discovers key phenomena—such as the career advancement of lower income youths and their ability (or inability) to break neighborhood ties.

The study has been highly regarded despite its being a single-case study, covering one neighborhood ("Cornerville") and a time period now more than 50 years old. The value of the book is, paradoxically, its generalizability to issues on individual performance, group structure, and the social structure of neighborhoods. Later investigators have repeatedly found remnants of Cornerville in their work, even though they have studied different neighborhoods and different time periods.

explanatory. There may be exploratory case studies, descriptive case studies, or explanatory case studies (Yin, 1981a, 1981b). There also may be exploratory experiments, descriptive experiments, and explanatory experiments. What distinguishes the strategies is not this hierarchy but three other conditions, discussed below. Nevertheless, this does not imply that the boundaries between the strategies—or the occasions when each is to be used—are always clear and sharp. Even though each strategy has its distinctive characteristics, there are large areas of overlap among them (e.g., Sieber, 1973). The goal is to avoid gross misfits—that is, when you are planning to use one type of strategy but another is really more advantageous.

When to Use Each Strategy

The three conditions consist of (a) the type of research question posed, (b) the extent of control an investigator has over actual behavioral events, and (c) the degree of focus on contemporary as opposed to historical events. Figure 1.1 displays these three conditions and shows how each is related to five major research strategies in the social sciences: experiments, surveys, archival analysis, histories, and case studies. The importance of each condition, in distinguishing among the five strategies, is discussed below.

BOX 2
An Explanatory Case Study

Even a single-case study can often be used to pursue an explanatory, and not merely exploratory (or descriptive), purpose. The analyst's objective should be to pose competing explanations for the same set of events and to indicate how such explanations may apply to other situations.

This strategy was followed by Graham Allison in *Essence of Decision: Explaining the Cuban Missile Crisis* (1971). The single case is the confrontation between the United States and the Soviet Union over the placement of offensive missiles in Cuba. Allison posits three competing theories or models to explain the course of events, including answers to three key questions: why the Soviet Union placed offensive (and not merely defensive) missiles in Cuba in the first place, why the United States responded to the missile deployment with a blockade (and not an air strike or invasion), and why the Soviet Union eventually withdrew the missiles. By comparing each theory with the actual course of events, Allison develops the best explanation for this type of crisis.

Allison suggests that this explanation is applicable to other situations, thereby extending the usefulness of his single-case study. Thus Allison cites the U.S. involvement in Vietnam, nuclear confrontation more generally, and the termination of wars by nations as other situations for which the theory can offer useful explanation.

Types of research questions (Figure 1.1, column 1). The first condition covers your research question(s) (Hedrick, Bickman, & Rog, 1993). A basic categorization scheme for the types of questions is the familiar series: "who," "what," "where," "how," and "why."

If research questions focus mainly on "what" questions, either of two possibilities arises. First, some types of "what" questions are exploratory, such as this one: "What are the ways of making schools effective?" This type of question is a justifiable rationale for conducting an exploratory study, the goal being to develop pertinent hypotheses and propositions for further inquiry. However, as an exploratory study, any of the five research strategies can be used—for example, an exploratory survey, an exploratory experiment, or an exploratory case study. The second type of "what" question is actually a form of a "how many" or "how much" line of inquiry—for example, "What have been the outcomes from a particular managerial reorganization?" Identifying such outcomes is more likely to favor survey or archival strategies than others. For example, a survey can be readily designed to enumerate the

strategy	form of research question	requires control over behavioral events?	focuses on contemporary events?
experiment	how, why	yes	yes
survey	who, what, where, how many, how much	no	yes
archival analysis	who, what, where, how many, how much	no	yes/no
history	how, why	no	no
case study	how, why	no	yes

Figure 1.1. Relevant Situations for Different Research Strategies
SOURCE: COSMOS Corporation.

"whats," whereas a case study would not be an advantageous strategy in this situation.

Similarly, like this second type of "what" question, "who" and "where" questions (or their derivatives—"how many" and "how much") are likely to favor survey strategies or the analysis of archival records, as in economic research. These strategies are advantageous when the research goal is to describe the incidence or prevalence of a phenomenon or when it is to be *predictive* about certain outcomes. The investigation of prevalent political attitudes (in which a survey or a poll might be the favored strategy) or of the spread of a disease like AIDS (in which an analysis of health statistics might be the favored strategy) would be typical examples.

In contrast, "how" and "why" questions are more *explanatory* and likely to lead to the use of case studies, histories, and experiments as the preferred research strategies. This is because such questions deal with operational links needing to be traced over time, rather than mere frequencies or incidence. Thus, if you wanted to know how a community successfully thwarted a proposed highway (see Lupo et al., 1971), you would be less likely to rely on a survey or an examination of archival records and might be better off doing a history or a case study. Similarly, if you wanted to know why bystanders fail to report emergencies under certain conditions, you could design and conduct a series of experiments (see Latané & Darley, 1969).

Let us take two more examples. If you were studying "who" participated in riots, and "how much" damage had been done, you might survey residents, examine business records (an archival analysis), or conduct a "windshield survey" of the riot area. In contrast, if you wanted to know "why" riots occurred, you would have to draw upon a wider array of documentary information, in addition to conducting interviews; if you focused on the "why" question in more than one city, you would probably be doing a multiple-case study.

Similarly, if you wanted to know "what" the outcomes of a new governmental program had been, you could answer this frequency question by doing a survey or by examining economic data, depending upon the type of program involved. Thus consider such questions as these: How many clients did the program serve? What kinds of benefits were received? How often were different benefits produced? These could all be answered without doing a case study. But if you needed to know "how" or "why" the program had worked (or not), you would lean toward either a case study or a field experiment.

Some "how" and "why" questions are ambivalent and need clarification. "How" and "why" Bill Clinton got elected in 1992 can be studied by either a survey or a case study. The survey might examine voting patterns, showing that voters for Ross Perot drew largely from supporters of then President Bush, and this could satisfactorily address the how and why questions. In contrast, the case study might examine how Clinton conducted his campaign to achieve the necessary nomination and to manipulate public opinion in his favor. The study would cover the potentially helpful role of the weak U.S. economy in denying support for the Bush-Quayle ticket as incumbents. This approach also would be an acceptable way of addressing the "how" and "why" questions but would be different than the survey study.

To summarize, the first and most important condition for differentiating among the various research strategies is to identify the type of research question being asked. In general, "what" questions may either be exploratory (in which case any of the strategies could be used) or about prevalence (in which surveys or the analysis of archival records would be favored). "How" and "why" questions are likely to favor the use of case studies, experiments, or histories.

Defining the research questions is probably the most important step to be taken in a research study, so patience and sufficient time should be allowed for this task. The key is to understand that research questions have both *substance*—for example, What is my study about?—and *form*—for example, Am I asking a "who," "what," "where," "why," or "how" question? Others have focused on some of the substantively important issues (see Campbell,

Daft, & Hulin, 1982); the point of the preceding discussion is that the form of the question provides an important clue regarding the appropriate research strategy to be used. Remember, too, the large areas of overlap among the strategies, so that, for some questions, a choice among strategies might actually exist. Remember, finally, that you may be predisposed to pursue a particular strategy regardless of the study question. If so, be sure to create the form of the study question best matching the strategy you were inclined to pursue in the first place.

Extent of control over behavioral events (Figure 1.1, column 2) and degree of focus on contemporary as opposed to historical events (Figure 1.1, column 3). Assuming that "how" and "why" questions are to be the focus of study, a further distinction among history, case study, and experiment is the extent of the investigator's control over and access to actual behavioral events. Histories are the preferred strategy when there is virtually no access or control. Thus the distinctive contribution of the historical method is in dealing with the "dead" past—that is, when no relevant persons are alive to report, even retrospectively, what occurred, and when an investigator must rely on primary documents, secondary documents, and cultural and physical artifacts as the main sources of evidence. Histories can, of course, be done about contemporary events; in this situation, the strategy begins to overlap with that of the case study.

The case study is preferred in examining contemporary events, but when the relevant behaviors cannot be manipulated. The case study relies on many of the same techniques as a history, but it adds two sources of evidence not usually included in the historian's repertoire: direct observation and systematic interviewing. Again, although case studies and histories can overlap, the case study's unique strength is its ability to deal with a full variety of evidence—documents, artifacts, interviews, and observations—beyond what might be available in the conventional historical study. Moreover, in some situations, such as participant-observation, informal manipulation can occur.

Finally, experiments are done when an investigator can manipulate behavior directly, precisely, and systematically. This can occur in a laboratory setting, in which an experiment may focus on one or two isolated variables (and presumes that the laboratory environment can "control" for all the remaining variables beyond the scope of interest), or it can be done in a field setting, where the term *social experiment* has emerged to cover research in which investigators "treat" whole groups of people in different ways, such as providing them with different kinds of vouchers (Boruch, forthcoming). Again, the methods overlap. The full range of experimental science also

includes those situations in which the experimenter cannot manipulate behavior (see Blalock, 1961; Campbell & Stanley, 1966; Cook & Campbell, 1979) but in which the logic of experimental design may still be applied. These situations have been commonly regarded as "quasi-experimental" situations. The quasi-experimental approach can even be used in a historical setting, in which, for instance, an investigator may be interested in studying race riots or lynchings (see Spilerman, 1971) and may use a quasi-experimental design because no control over the behavioral event was possible.

Summary. We can identify some situations in which all research strategies might be relevant (such as exploratory research), and other situations in which two strategies might be considered equally attractive (such as how and why Clinton got elected). We also can use more than one strategy in any given study (for example, a survey within a case study or a case study within a survey). To this extent, the various strategies are not mutually exclusive. But we can also identify some situations in which a specific strategy has a distinct advantage. For the *case study*, this is when

- a "how" or "why" question is being asked about a contemporary set of events over which the investigator has little or no control.

To determine the questions that are most significant for a topic, and to gain some precision in formulating these questions, requires much preparation. One way is to review the literature on the topic (Cooper, 1984). Note that such a literature review is therefore a means to an end, and not—as most students think—an end in itself. Budding investigators think that the purpose of a literature review is to determine the *answers* about what is known on a topic; in contrast, experienced investigators review previous research to develop sharper and more insightful *questions* about the topic.

Traditional Prejudices Against the Case Study Strategy

Although the case study is a distinctive form of empirical inquiry, many research investigators nevertheless have disdain for the strategy. In other words, as a research endeavor, case studies have been viewed as a less desirable form of inquiry than either experiments or surveys. Why is this?

Perhaps the greatest concern has been over the lack of rigor of case study research. Too many times, the case study investigator has been sloppy and has allowed equivocal evidence or biased views to influence the direction of the findings and conclusions.

The possibility also exists that people have confused case study teaching with case study research. In teaching, case study materials may be deliberately altered to demonstrate a particular point more effectively. In research, any such step would be strictly forbidden. Every case study investigator must work hard to report all evidence fairly, and this book will help him or her to do so. What is often forgotten is that bias also can enter into the conduct of experiments (see Rosenthal, 1966) and the use of other research strategies, such as designing questionnaires for surveys (Sudman & Bradburn, 1982) or conducting historical research (Gottschalk, 1968). The problems are not different, but in case study research, they may have been more frequently encountered and less frequently overcome.

A second common concern about case studies is that they provide little basis for scientific generalization. "How can you generalize from a single case?" is a frequently heard question. The answer is not a simple one (Kennedy, 1976). However, consider for the moment that the same question had been asked about an experiment: "How can you generalize from a single experiment?" In fact, scientific facts are rarely based on single experiments; they are usually based on a multiple set of experiments, which have replicated the same phenomenon under different conditions. The same approach can be used with multiple-case studies but requires a different concept of the appropriate research designs; this is discussed in detail in Chapter 2. The short answer is that case studies, like experiments, are generalizable to theoretical propositions and not to populations or universes. In this sense, the case study, like the experiment, does not represent a "sample," and the investigator's goal is to expand and generalize theories (analytic generalization) and not to enumerate frequencies (statistical generalization). Or, as three notable social scientists describe in their *single* case study, the goal is to do a "generalizing" and not a "particularizing" analysis (Lipset, Trow, & Coleman, 1956, pp. 419-420).

A third frequent complaint about case studies is that they take too long, and they result in massive, unreadable documents. This complaint may be appropriate, given the way case studies have been done in the past (e.g., Feagin, Orum, & Sjoberg, 1991), but this is not necessarily the way case studies must be done in the future. Chapter 6 discusses alternative ways of writing the case study—including ones in which the traditional, lengthy narrative can be avoided altogether. Nor need case studies take a long time. This incorrectly confuses the case study strategy with a specific method of data collection, such as ethnography or participant-observation. Ethnographies usually require long periods of time in the "field" and emphasize detailed, observational evidence. Participant-observation may not require the same length of time but still assumes a hefty investment of field efforts. In contrast, case studies are a form of inquiry that does *not* depend solely on ethnographic or participant-observer data. One could even do a valid and high-quality case study without leaving the library and the telephone, depending upon the topic being studied.

Despite the fact that these common concerns can be allayed, as above, one major lesson is still that good case studies are very difficult to do. The problem is that we have little way of screening or testing for an investigator's ability to do good case studies. People know when they cannot play music; they also know when they cannot do mathematics; and they can be tested for other skills, such as by the bar examination in law. Somehow, the skills for doing good case studies have not yet been defined, and as a result,

> most people feel that they can prepare a case study, and nearly all of us believe we can understand one. Since neither view is well founded, the case study receives a good deal of approbation it does not deserve. (Hoaglin, Light, McPeek, Mosteller, & Stoto, 1982, p. 134)

This quotation is from a book by five prominent *statisticians*. Surprisingly, even from another field, they recognize the challenge of doing good case studies.

DIFFERENT TYPES OF CASE STUDIES, BUT A COMMON DEFINITION

The discussion has progressed without a formal definition of case studies. Moreover, commonly asked questions about case studies have still been unanswered. For example, is it still a case study when more than one case is included in the same study? Do case studies preclude the use of quantitative evidence? Can case studies be used to do evaluations? Can case studies include journalistic accounts? Let us now attempt to define the case study strategy and answer these questions.

Definition of the Case Study as a Research Strategy

The most frequently encountered definitions of case studies have merely repeated the types of topics to which case studies have been applied. For example, in the words of one observer,

> the essence of a case study, the central tendency among all types of case study, is that it tries to illuminate a *decision* or set of decisions: why they were taken, how they were implemented, and with what result. (Schramm, 1971, emphasis added)

This definition thus cites the topic of "decisions" as the major focus of case studies. Similarly, other topics have been listed, including "individuals," "organizations," "processes," "programs," "neighborhoods," "institutions," and even "events." However, citing the topic is surely insufficient for establishing the needed definition.

Alternatively, most social science textbooks have failed to consider the case study a formal research strategy at all (the major exception is the book by five statisticians from Harvard University—Hoaglin et al., 1982). As discussed earlier, one common flaw was to consider the case study as the exploratory stage of some other type of research strategy, and the case study itself was only mentioned in a line or two of text.

Another common flaw has been to confuse case studies with ethnographies (Fetterman, 1989) or with participant-observation (Jorgensen, 1989), so that a textbook's presumed discussion of case studies was in reality a description either of the ethnographic method or of participant-observation as a data collection technique. The most popular contemporary texts (e.g., Kidder & Judd, 1986; Nachmias & Nachmias, 1992), in fact, still cover "fieldwork" only as a data collection technique and omit any further discussion of case studies.

In a historical overview of the case study in American methodological thought, Jennifer Platt (1992a) explains the reasons for these treatments. She traces the practice of doing case studies back to the conduct of life histories, the work of the Chicago school of sociology, and casework in social work. She then shows how "participant-observation" emerged as a data collection technique, leaving the further definition of any distinctive case study strategy in suspension. Finally, she explains how the first edition of this book (1984) definitively dissociated the case study strategy from the limited perspective of doing participant-observation (or any type of fieldwork). The case study strategy, in her words, begins with "a logic of design . . . a strategy to be preferred when circumstances and research problems are appropriate rather than an ideological commitment to be followed whatever the circumstances" (Platt, 1992a, p. 46).

And just what is this logic of design? The technically critical features had been worked out prior to the first edition of this book (Yin, 1981a, 1981b) but now may be restated in two ways. First, the technical definition begins with the scope of a case study:

1. A case study is an empirical inquiry that

- investigates a contemporary phenomenon within its real-life context, especially when
- the boundaries between phenomenon and context are not clearly evident.

In other words, you would use the case study method because you deliberately wanted to cover contextual conditions—believing that they might be highly pertinent to your phenomenon of study. This first part of our logic of design therefore helps us to understand case studies by continuing to distinguish them from the other research strategies that have been discussed.

An experiment, for instance, deliberately divorces a phenomenon from its context, so that attention can be focused on only a few variables (typically, the context is "controlled" by the laboratory environment). A history, by comparison, does deal with the entangled situation between phenomenon and context, but usually with *non*contemporary events. Finally, surveys can try to deal with phenomenon and context, but their ability to investigate the context is extremely limited. The survey designer, for instance, constantly struggles to limit the number of variables to be analyzed (and hence the number of questions that can be asked) to fall safely within the number of respondents that can be surveyed.

Second, because phenomenon and context are not always distinguishable in real-life situations, a whole set of other technical characteristics, including data collection and data analysis strategies, now become the second part of our technical definition:

2. The case study inquiry

- copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result
- relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result
- benefits from the prior development of theoretical propositions to guide data collection and analysis.

In other words, the case study as a research strategy comprises an all-encompassing method—with the logic of design incorporating specific approaches to data collection and to data analysis. In this sense, the case study is not either a data collection tactic or merely a design feature alone (Stoecker, 1991) but a comprehensive research strategy.[1] How the strategy is defined and implemented is the topic of this entire book.

Certain other features of the case study strategy are not critical for defining the strategy but may be considered variations within case study research and also provide answers to common questions.

Variations Within Case Studies as a Research Strategy

Yes, case study research can include both single- and multiple-case studies. Though some fields, such as political science and public administration, have tried to delineate sharply between these two approaches (and have used such terms as the *comparative case method* as a distinctive form of multiple-case studies; see Agranoff & Radin, 1991; George, 1979; Lijphart, 1975), single- and multiple-case studies are in reality but two variants of case study designs (see Chapter 2 for more).

And, yes, case studies can include, and even be limited to, quantitative evidence. In fact, the contrast between quantitative and qualitative evidence does not distinguish the various research strategies. Note that, as analogous examples, some experiments (such as studies of psychophysical perceptions) and some survey questions (such as those seeking categorical rather than numerical responses) rely on qualitative, and not quantitative, evidence. Likewise, historical research can include enormous amounts of quantitative evidence.

As a related but important note, the case study strategy should not be confused with "qualitative research" (see Schwartz & Jacobs, 1979; Strauss & Corbin, 1990; Van Maanen, 1988; Van Maanen, Dabbs, & Faulkner, 1982). Some qualitative research follows ethnographic methods and seeks to satisfy two conditions: (a) the use of close-up, detailed observation of the natural world by the investigator and (b) the attempt to avoid prior commitment to any theoretical model (Jacob, 1987, 1989; Lincoln & Guba, 1986; Stake, 1983; Van Maanen et al., 1982, p. 16). However, ethnographic research does not always produce case studies (for example, see the brief ethnographies in G. Jacobs, 1970), nor are case studies limited to these two conditions. Instead, case studies can be based on any mix of quantitative and qualitative evidence. In addition, case studies need not always include direct, detailed observations as a source of evidence.

As a further note, some investigators distinguish between quantitative research and qualitative research—not on the basis of the type of evidence but on the basis of wholly different philosophical beliefs (e.g., Guba & Lincoln, 1989; Lincoln, 1991; Sechrest, 1991; Smith & Heshusius, 1986). These distinctions have produced a sharp debate within the field of evaluation research. Although some believe that these philosophical beliefs are irreconcilable, the counterargument can still be posed—that regardless of whether one favors qualitative or quantitative research, there is a strong and essential common ground between the two (Yin, 1994).

And, yes, case studies have a distinctive place in evaluation research (see Cronbach et al., 1980; Guba & Lincoln, 1981; Patton, 1980; U.S. General Accounting Office, 1990; Yin, 1993, chap. 4). There are at least five different applications. The most important is to *explain* the causal links in real-life interventions that are too complex for the survey or experimental strategies. In evaluation language, the explanations would link program implementation with program effects (U.S. General Accounting Office, 1990). A second application is to *describe* an intervention and the real-life context in which it occurred. Third, case studies can *illustrate* certain topics within an evaluation, again in a descriptive mode—even from a journalistic perspective. Fourth, the case study strategy may be used to *explore* those situations in which the intervention being evaluated has no clear, single set of outcomes. Fifth, the case study may be a "*meta-evaluation*"—a study of an evaluation study (N. Smith, 1990; Stake, 1986). Whatever the application, one constant theme is that program sponsors—rather than research investigators alone—may have the prominent role in defining the evaluation questions and relevant data categories (U.S. General Accounting Office, 1990).

And, finally, yes, certain journalistic efforts can qualify as case studies. Actually, one of the best written and most interesting case studies is about the Watergate scandal, by two reporters from *The Washington Post* (see BOX 3).

SUMMARY

This chapter has introduced the importance of the case study as a research strategy. The case study, like other research strategies, is a way of investigating an empirical topic by following a set of prespecified procedures. These procedures will largely dominate the remainder of this book.

The chapter also has attempted to distinguish the case study from alternative research strategies in social science, indicating the situations in which doing a single- or multiple-case study may be preferred, for instance, to doing a survey. Some situations may have no clearly preferred strategy, as the strengths and weaknesses of the various strategies may overlap. The basic approach, however, is to consider all the strategies in a pluralistic fashion—as part of a repertoire for doing social science research from which the investigator may draw according to a given situation.

BOX 3
A Journalistic Case Study

Although public memory of President Richard M. Nixon's resignation has receded, Bernstein and Woodward's *All the President's Men* (1974) remains a fascinating account of the Watergate scandal. The book is dramatic and suspenseful, relies on solid journalistic methods, and serendipitously represents a common design for case studies.

The "case," in this book, is not the Watergate burglary itself, or even the Nixon administration more generally. Rather, the case is the "coverup," a complex set of events that occurred in the aftermath of the burglary. Bernstein and Woodward continually confront the reader with two "how" and "why" questions: How did the coverup occur, and why did it occur? Neither is answered easily, and the book's appeal lies in its piecing together of fact after fact, each piece adding up curiously and then potently to an explanation for the coverup.

Establishing the how and why of a complex human situation is a classic example of the use of case studies, whether done by journalists or social scientists. If the case involves a significant public event and an appealing explanation, the ingredients may add up, as in *All the President's Men*, to a best-seller.

Finally, the chapter has discussed some of the major criticisms of case study research and has suggested that these criticisms are misdirected. However, we must all work hard to overcome the problems of doing case study research, including the recognition that some of us were not meant, by skill or disposition, to do such research in the first place. Case study research is remarkably hard, even though case studies have traditionally been considered to be "soft" research. Paradoxically, the "softer" a research strategy, the harder it is to do.

EXERCISES

1. *Defining a case study question.* Develop a question that would be the rationale for a case study you might conduct. Instead of doing a case study, now imagine that you could only do a history, a survey, or an experiment (but not a case study) in order to answer this question. What aspects of the question, if any, could not be answered through these other research strategies? What would be the distinctive advantage of doing a case study to answer this question?

2. *Defining "significant" case study questions.* Name a topic you think is worthy of making the subject of a case study. Identify the three major questions your case study would try to answer. Now assume that you were actually able to answer these questions with sufficient evidence (i.e., that you had successfully conducted your case study). How would you justify, to a colleague, the significance of your findings? Would you have advanced some major theory? Would you have discovered something rare? (If you are unimpressed by your answers, perhaps you should consider redefining the major questions of your case.)

3. *Identifying "significant" questions in other research strategies.* Locate a research study based solely on the use of survey, historical, or experimental (but not case study) methods. Describe the ways in which the findings of this study are significant. Does it advance some major theory? Has it discovered something rare?

4. *Examining case studies used for teaching purposes.* Obtain a copy of a case study designed for teaching purposes (e.g., a case in a textbook used in a business school course). Identify the specific ways in which this type of "teaching" case is different than research case studies. Does the teaching case cite primary documents, contain evidence, or display data? Does the teaching case have a conclusion? What appears to be the main objective of the teaching case?

5. *Defining different types of case studies used for research purposes.* Define the three types of case studies used for research (but not teaching) purposes: (a) explanatory or causal case studies, (b) descriptive case studies, and (c) exploratory case studies. Compare the situations in which these different types of case studies would be most applicable, and then name a case study you would like to conduct. Would it be explanatory, descriptive, or exploratory? Why?

NOTE

1. Robert Stake (1994) has yet another approach for defining case studies. He considers them not to be "a methodological choice but a choice of object to be studied." Further, the object must be a "functioning specific" (such as a person or classroom) but not a generality (such as a policy). This definition is too broad. Every study of entities qualifying as objects (e.g., people, organizations, and countries) would then be a case study, regardless of the methodology used (e.g., psychological experiment, management survey, economic analysis).

2

Designing Case Studies

A research design is the logic that links the data to be collected (and the conclusions to be drawn) to the initial questions of a study. Every empirical study has an implicit, if not explicit, research design.

For case studies, four major types of designs are relevant, following a 2×2 matrix. The first pair of categories consists of single-case and multiple-case designs. The second pair, which can occur in combination with either of the first pair, is based on the unit or units of analysis to be covered—and distinguishes between holistic and embedded designs.

The case study investigator also must maximize four aspects of the quality of any design: (a) construct validity, (b) internal validity (for explanatory or causal case studies only), (c) external validity, and (d) reliability. How the investigator should deal with these four aspects of quality control is summarized in Chapter 2 but also is a major theme throughout the remainder of the book.

GENERAL APPROACH TO DESIGNING CASE STUDIES

In identifying the research strategy for your research project, Chapter 1 has shown when you should select the case study strategy, as opposed to other strategies. The next task is to design your case study. For this purpose, as in designing any other type of research investigation, a plan, or *research design*, is needed.

The development of this research design is a difficult part of doing case studies. Unlike other research strategies, a comprehensive "catalog" of research designs for case studies has yet to be developed. There are no textbooks like those in the biological and psychological sciences, covering such design considerations as the assignment of subjects to different "groups," the selection of different stimuli or experimental conditions, or the identification of various response measures (see Cochran & Cox, 1957; Fisher, 1935, cited in Cochran & Cox, 1957; Sidowski, 1966). In a laboratory experiment, each of these choices reflects an important logical connection to the issues being studied. Similarly, there are not even textbooks like the well-known volumes

by Campbell and Stanley (1966) or by Cook and Campbell (1979), which summarize the various research designs for quasi-experimental situations. Nor have there emerged any common designs—for example, "panel" studies—such as those now recognized in doing survey research (see Kidder & Judd, 1986, chap. 6).

One pitfall to be avoided, however, is to consider case study designs to be a subset or variant of the research designs used for other strategies, such as experiments. For the longest time, scholars incorrectly thought that the case study was but one type of quasi-experimental design (the one-shot posttest-only design). This misperception has finally been corrected, with the following statement appearing in a revision on quasi-experimental designs: "Certainly the case study as normally practiced should not be demeaned by identification with the one-group post-test-only design" (Cook & Campbell, 1979, p. 96).

In other words, the one-shot, posttest-only design as a quasi-experimental design still may be considered flawed, but the case study has now been recognized as something different. In fact, the case study is a separate research strategy that has its own research designs.

Unfortunately, case study research designs have not been codified. The following chapter therefore expands on the new methodological ground broken by the first edition of this book and describes a basic set of research designs for doing single- and multiple-case studies. Although these designs will need to be continually modified and improved in the future, in their present form they will nevertheless help you to design more rigorous and methodologically sound case studies.

Definition of Research Designs

Every type of empirical research has an implicit, if not explicit, research design. In the most elementary sense, the design is the logical sequence that connects the empirical data to a study's initial research questions and, ultimately, to its conclusions. Colloquially, a research design is *an action plan for getting from here to there*, where *here* may be defined as the initial set of questions to be answered, and *there* is some set of conclusions (answers) about these questions. Between "here" and "there" may be found a number of major steps, including the collection and analysis of relevant data. As a summary definition, another textbook has described a research design as a plan that

> guides the investigator in the process of collecting, analyzing, and interpreting observations. It is a *logical model of proof* that allows the researcher to draw

> inferences concerning causal relations among the variables under investigation. The research design also defines the domain of generalizability, that is, whether the obtained interpretations can be generalized to a larger population or to different situations. (Nachmias & Nachmias, 1992, pp. 77-78, emphasis added)

Another way of thinking about a research design is as a "blueprint" of research, dealing with at least four problems: what questions to study, what data are relevant, what data to collect, and how to analyze the results (see F. Borum, personal communication, Copenhagen Business School, Copenhagen, Denmark, 1991; Philliber, Schwab, & Samsloss, 1980).

Note that a research design is much more than a work plan. The main purpose of the design is to help to avoid the situation in which the evidence does not address the initial research questions. In this sense, a research design deals with a *logical* problem and not a *logistical* problem. As a simple example, suppose you want to study a single organization. Your research questions, however, have to do with the organization's relationships with other organizations—their competitive or collaborative nature, for example. Such questions can be answered only if you collect information directly from the other organizations and not merely from the one you started with. If you complete your study by examining only one organization, you cannot draw accurate conclusions about interorganizational partnerships. This is a flaw in your research design, not in your work plan. The outcome could have been avoided if you had developed an appropriate research design in the first place.

Components of Research Designs

For case studies, five components of a research design are especially important:

1. a study's questions,
2. its propositions, if any,
3. its unit(s) of analysis,
4. the logic linking the data to the propositions, and
5. the criteria for interpreting the findings.

Study questions. This first component has already been described in Chapter 1. Although the substance of your questions will vary, Chapter 1 suggested that the *form* of the question—in terms of "who," "what," "where," "how," and "why"—provides an important clue regarding the most relevant research strategy to be used. The case study strategy is most likely to

be appropriate for "how" and "why" questions, so your initial task is to clarify precisely the nature of your study questions in this regard.

Study propositions. As for the second component, each proposition directs attention to something that should be examined within the scope of the study. For instance, assume that your research, on the topic of interorganizational partnerships, began with the question: How and why do organizations collaborate with one another to provide joint services (for example, a manufacturer and a retail store collaborating to sell certain computer products)? These "how" and "why" questions, capturing what you are really interested in answering, led you to the case study as the appropriate strategy in the first place. Nevertheless, these "how" and "why" questions do not point to what you should study. Only if you are forced to state some propositions will you move in the right direction. For instance, you might think that organizations collaborate because they derive mutual benefits. This proposition, in addition to reflecting an important theoretical issue (that other incentives for collaboration do not exist or are unimportant), also begins to tell you where to look for relevant evidence (to define and ascertain the extent of specific benefits to each organization).

At the same time, some studies may have a legitimate reason for not having any propositions. This is the condition—which exists in experiments, surveys, and the other research strategies alike—in which a topic is the subject of "exploration." Every exploration, however, should still have some purpose. Instead of stating propositions, the design for an exploratory study should state a purpose, as well as the criteria by which an exploration will be judged successful. Consider the analogy in BOX 4 for exploratory case studies. Can you imagine how you would ask for support from Queen Isabella to do your exploratory study?

Unit of analysis. This third component is related to the fundamental problem of defining what the "case" is—a problem that has plagued many investigators at the outset of case studies. For instance, in the classic case study, a "case" may be an individual. Jennifer Platt (1992a, 1992b) has noted how the early case studies in the Chicago school of sociology were life histories of such roles as juvenile delinquents or derelict men. You also can imagine case studies of clinical patients, of exemplary students, or of certain types of leaders. In each situation, an individual person is the case being studied, and the individual is the primary unit of analysis. Information about each relevant individual would be collected, and several such individuals or "cases" might be included in a multiple-case study. Propositions would still be needed to help identify the relevant information about this individual or

BOX 4
"Exploration" as an Analogy for an Exploratory Case Study

When Christopher Columbus went to Queen Isabella to ask for support for his "exploration" of the New World, he had to have some reasons for asking for three ships (Why not one? Why not five?), and he had some rationale for going westward (Why not south? Why not south and then east?). He also had some (mistaken) criteria for recognizing the Indies when he actually encountered them. In short, his exploration began with some rationale and direction, even if his initial assumptions might later have been proved wrong (Wilford, 1992). This same degree of rationale and direction should underlie even an exploratory case study.

individuals. Without such propositions, an investigator might be tempted to collect "everything," which is impossible to do. For example, the propositions in studying these individuals might involve the influence of early childhood or the role of peer relationships. Such topics already represent a vast narrowing of the relevant data. The more a study contains specific propositions, the more it will stay within feasible limits.

Of course, the "case" also can be some event or entity that is less well defined than a single individual. Case studies have been done about decisions, about programs, about the implementation process, and about organizational change. Feagin, Orum, & Sjoberg (1991) contains some classic examples of these single cases in sociology and political science. Beware of these types of topics—none is easily defined in terms of the beginning or end points of the "case." For example, a case study of a specific program may reveal (a) variations in program definition, depending upon the perspective of different actors, and (b) program components that existed prior to the formal designation of the program. Any case study of such a program would therefore have to confront these conditions in delineating the unit of analysis.

As a general guide, the definition of the unit of analysis (and therefore of the case) is related to the way the initial research questions have been defined. Suppose, for example, you want to study the role of the United States in the world economy. Peter Drucker (1986) has written a provocative essay about fundamental changes in the world economy, including the importance of "capital movements" independent of the flow of goods and services. The unit of analysis for your case study might be a country's economy, an industry in

BOX 5a
What Is the Unit of Analysis?

The Soul of a New Machine (1981) was a Pulitzer prize-winning book by Tracy Kidder. The book, also a best-seller, is about the development of a new minicomputer produced by Data General Corporation, intended to compete directly with one produced by Digital Equipment Corporation.

This easy-to-read book describes how Data General's engineering team invented and developed the new computer. The book begins with the initial conceptualization of the computer and ends when the engineering team relinquishes control of the machine to Data General's marketing staff.

The book is an excellent example of a case study. However, the book also illustrates a fundamental problem in doing case studies—that of defining the *unit of analysis*. Is the case study about the minicomputer, or is it about the dynamics of a small group—the engineering team? The answer is critical if we want to understand how the case study relates to a broader body of knowledge—that is, whether to generalize to a technology topic or to a group dynamics topic. Because the book is not an academic study, it does not need to, nor does it, provide an answer.

the world marketplace, an economic policy, or the trade or capital flow between two countries. Each unit of analysis would call for a slightly different research design and data collection strategy. Selection of the appropriate unit of analysis results from your accurately specifying the primary research questions. If your questions do not lead to the favoring of one unit of analysis over another, your questions are probably either too vague or too numerous—and you may have trouble conducting your case study.

Sometimes, the unit of analysis may have been defined one way, even though the phenomenon being studied calls for a different definition. Most frequently, investigators have confused case studies of neighborhoods with case studies of small groups (for another example, confusing an innovation with a small group in organizational studies, see BOX 5a). How a general *area* such as a neighborhood copes with racial transition, upgrading, and other phenomena can be quite different than how a small *group* copes with these same phenomena. *Street Corner Society* (Whyte, 1943/1955—also see BOX 1 in Chapter 1 of this book) and *Tally's Corner* (Liebow, 1967—also see BOX 9, this chapter), for instance, have often been mistaken for being case studies of neighborhoods when in fact they are case studies of small groups (note that in neither book is the neighborhood geography described, even though

BOX 5b
A Clearer Choice Among Units of Analysis

Ira Magaziner and Mark Patinkin's book *The Silent War: Inside the Global Business Battles Shaping America's Future* (1989) presents nine case studies. Each case study helps the reader to understand a real-life situation of international economic competition.

Two of the cases appear similar but in fact have different main units of analysis. One case, about the Korean firm Samsung, is a case study of the critical policies that make the firm competitive. Understanding Korean economic development is part of the context, and the case study also contains an embedded unit—Samsung's development of the microwave oven as an illustrative product. The other case, about the development of an Apple computer factory in Singapore, is in fact a case study of Singapore's critical policies that make the country competitive. The Apple computer factory experience—an embedded unit of analysis—is actually an illustrative example of how national policies affect foreign investments.

These two cases show how the definition of the main and embedded units of analyses, as well as the definition of the contextual events surrounding these units, depends on the level of inquiry. The main unit of analysis is likely to be at the level being addressed by the main study questions.

the small groups lived in a small area with clear neighborhood implications). BOX 5b, however, presents a good example of how units of analyses can be defined in a more discriminating manner—in the field of world trade.

Most investigators will encounter this type of confusion in defining the unit of analysis. To reduce the confusion, one good practice is to discuss the potential case with a colleague. Try to explain to that person what questions you are trying to answer and why you have chosen a specific case or group of cases as a way of answering those questions. This may help you to avoid incorrectly identifying the unit of analysis.

Once the general definition of the case has been established, other clarifications in the unit of analysis become important. If the unit of analysis is a small group, for instance, the persons to be included within the group (the immediate topic of the case study) must be distinguished from those who are outside it (the context for the case study). Similarly, if the case is about services in a specific geographic area, decisions need to be made about public services whose district boundaries do not coincide with the area. Finally, for almost any topic that might be chosen, specific time boundaries are needed to define the beginning and end of the case. All of these types of questions need to be

considered and answered to define the unit of analysis and thereby to determine the limits of the data collection and analysis.

One final point needs to be made about defining the case and the unit of analysis, pertaining to the role of the available research literature. Most researchers will want to compare their findings with previous research; for this reason, the key definitions should not be idiosyncratic. Rather, each case study and unit of analysis either should be similar to those previously studied by others or should deviate in clear, operationally defined ways. In this manner, the previous literature therefore also can become a guide for defining the case and unit of analysis.

Linking data to propositions, and criteria for interpreting the findings. The fourth and fifth components have been the least well developed in case studies. These components represent the data analysis steps in case study research, and a research design should lay the foundations for this analysis.

Linking data to propositions can be done any number of ways, but none has become as precisely defined as the assignment of subjects and treatment conditions in psychological experiments (which is the way that hypotheses and data are connected in psychology). One promising approach for case studies is the idea of "pattern-matching" described by Donald Campbell (1975), whereby several pieces of information from the same case may be related to some theoretical proposition. In a related article on one type of pattern—a time-series pattern—Campbell (1969) illustrated this approach but without labeling it as such.

In his article, Campbell first showed how the annual number of traffic fatalities in Connecticut had seemed to decline after the passage of a new state law limiting the speed to 55 miles per hour. However, further examination of the fatality rate, over a number of years before and after the legal change, showed unsystematic fluctuation rather than any marked reduction. A simple eyeball test was all that was needed to show that the actual pattern *looked* unsystematic rather than following a downtrend (see Figure 2.1), and thus Campbell concluded that the speed limit had had no effect on the number of traffic fatalities.

What Campbell did was describe two potential patterns and then show that the data matched one better than the other. If the two potential patterns are considered rival propositions (an "effects" proposition and a "no effects" proposition, regarding the impact of the new speed limit law), the pattern-matching technique is a way of relating the data to the propositions, even though the entire study consists of only a single case (the state of Connecticut).

This article also illustrates the problems in dealing with the fifth component, *the criteria for interpreting a study's findings*. Campbell's data matched

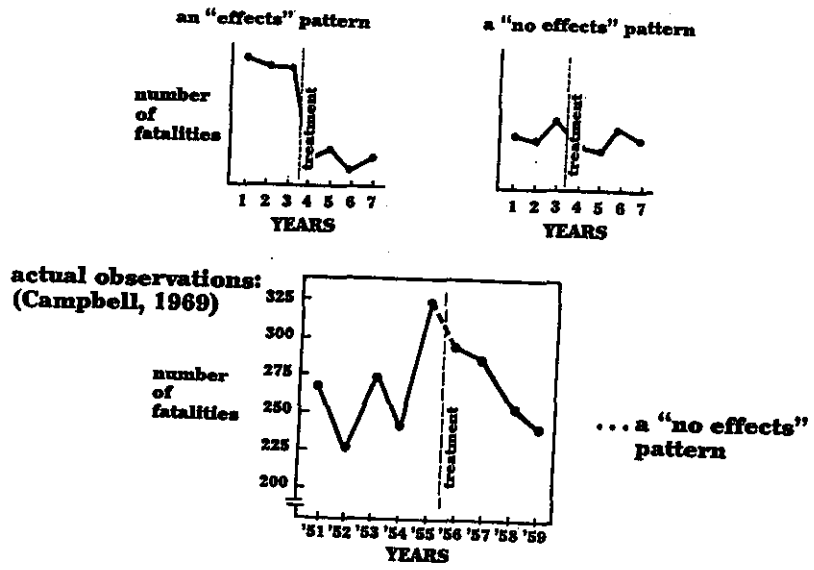

Figure 2.1. An Example of Pattern-Matching
SOURCE: COSMOS Corporation.

one pattern much better than they matched the other. But how close does a match have to be so as to be considered a match? Note that Campbell did not do any statistical test to make the comparison. Nor would a statistical test have been possible, because each data point in the pattern was a single number—the number of fatalities for that year—for which one could not calculate a variance and could not conduct any statistical test. Currently, there is no precise way of setting the criteria for interpreting these types of findings. One hopes that the different patterns are sufficiently contrasting that (as in Campbell's case) the findings can be interpreted in terms of comparing at least two rival propositions.

Summary. A research design should include five components. Although the current state of the art does not provide detailed guidance on the last two, the complete research design should not only indicate what data are to be collected—as indicated by (a) a study's questions, (b) its propositions, and (c) its units of analysis. The design also should tell you what is to be done after the data have been collected—as indicated by (d) the logic

linking the data to the propositions and (e) the criteria for interpreting the findings.

The Role of Theory in Design Work

Covering these preceding five components of research designs will effectively force you to begin constructing a preliminary theory related to your topic of study. This role of theory development, prior to the conduct of any data collection, is one point of difference between case studies and related methods such as ethnography (Lincoln & Guba, 1985, 1986; Van Maanen, 1988; Van Maanen et al., 1982) and "grounded theory" (Strauss & Corbin, 1990). Typically, these related methods deliberately avoid specifying any theoretical propositions at the outset of an inquiry. As a result, students wrongly think that by using the case study method, they can proceed quickly into the data collection phase of their work, and they have been encouraged to make their "field contacts" as quickly as possible. No guidance could be more misleading. Among other considerations, the relevant field contacts depend upon an understanding—or theory—of what is being studied.

Theory development. For case studies, theory development as part of the design phase is essential, whether the ensuing case study's purpose is to develop or to test theory. Using a case study on the implementation of a new management information system (MIS) as an example (Markus, 1983), the simplest ingredient of a theory is a statement such as the following:

> The case study will show why implementation only succeeded when the organization was able to re-structure itself, and not just overlay the new MIS on the old organizational structure. (Markus, 1983)

The statement presents the nutshell of a theory of MIS implementation—that is, that organizational restructuring is needed to make MIS implementation work.

Using the same case, an additional ingredient might be the following statement:

> The case study will also show why the simple replacement of key persons was not sufficient for successful implementation. (Markus, 1983)

This second statement presents the nutshell of a *rival* theory—that is, that MIS implementation fails because of the resistance to change on the part of

individual people, and that the replacement of such people is the only requirement for implementation to succeed.

You can see that, as these two initial ingredients are elaborated, the stated ideas will increasingly cover the questions, propositions, units of analysis, logic connecting data to propositions, and criteria for interpreting the findings—that is, the five components of the needed research design. In this sense, the complete research design embodies a "theory" of what is being studied. This theory should by no means be considered with the formality of grand theory in social science, nor are you being asked to be a masterful theoretician. Rather, the simple goal is to have a sufficient blueprint for your study, and this requires theoretical propositions. Then, the complete research design will provide surprisingly strong guidance in determining what data to collect and the strategies for analyzing the data. For this reason, theory development prior to the collection of any case study data is an essential step in doing case studies.

However, theory development takes time and can be difficult (Eisenhardt, 1989). For some topics, existing works may provide a rich theoretical framework for designing a specific case study. If you are interested in international economic development, for instance, Peter Drucker's "The Changed World Economy" (1986) is an exceptional source of theories and hypotheses. Drucker claims that the world economy has changed significantly from the past. He points to the "uncoupling" between the primary products (raw materials) economy and the industrial economy, a similar uncoupling between low labor costs and manufacturing production, and the uncoupling between financial markets and the real economy of goods and services. To test these propositions might require different studies, some focusing on the different uncouplings, others focusing on specific industries, and yet others explaining the plight of specific countries. Each different study would likely call for a different unit of analysis. Drucker's theoretical framework would provide guidance for designing these studies and even for collecting relevant data.

In other situations, the appropriate theory may be a descriptive theory (see BOX 6, and also BOX 1 for another example), and your concern should focus on such issues as (a) the purpose of the descriptive effort, (b) the full but realistic range of topics that might be considered a "complete" description of what is to be studied, and (c) the likely topic(s) that will be the essence of the description. Good answers to these questions, including the rationales underlying the answers, will help you go a long way toward developing the needed theoretical base—and research design—for your study.

For yet other topics, the existing knowledge base may be poor, and the available literature will provide no conceptual framework or hypotheses of note. Such a knowledge base does not lend itself to the development of good

BOX 6
Using a Metaphor to Develop Descriptive Theory

Whether four countries—the American colonies, Russia, England, and France—all underwent similar courses of events during their major political revolutions is the topic of Crane Brinton's famous historical study—*The Anatomy of a Revolution* (1938). Tracing and analyzing these events is done in a descriptive manner, as the author's purpose is not so much to explain the revolutions as to determine whether they followed similar courses.

The "cross-case" analysis reveals major similarities: All societies were on the upgrade, economically; there were bitter class antagonisms; the intellectuals deserted from positions of leadership; government machinery was inefficient; and the ruling class exhibited immoral, dissolute, or inept behavior (or all three). However, rather than relying solely on this "factors" approach to description, the author also develops the metaphor of a human body suffering from a fever as a way of describing the pattern of events over time. The author adeptly uses the cyclic pattern of fever and chills, rising to a critical point and followed by a false tranquility, to describe the ebb and flow of events in the four revolutions.

theoretical statements, and any new empirical study is likely to assume the characteristic of being an "exploratory" study. Nevertheless, as noted earlier with the illustrative case in BOX 4, even an exploratory case study should be preceded by statements about (a) what is to be explored, (b) the purpose of the exploration, and (c) the criteria by which the exploration will be judged successful.

Illustrative types of theories. In general, to overcome the barriers to theory development, you should try to prepare for your case study by doing such things as reviewing the literature related to what you would like to study (also see Cooper, 1984); discussing your topic and ideas with colleagues or teachers; and asking yourself challenging questions about what you are studying, why you are proposing to do the study, and what you hope to learn as a result of the study.

As a further reminder, you should be aware of the full range of theories that might be relevant to your study. For instance, note that the MIS example illustrates MIS "implementation" theory, and that this is but one type of theory that can be the subject of study. Other types of theories for you to consider include the following:

- Individual theories—for example, theories of individual development, cognitive behavior, personality, learning and disability, individual perception, and interpersonal interactions
- Group theories—for example, theories of family functioning, informal groups, work teams, supervisory-employee coordination, and interpersonal networks
- Organizational theories—for example, theories of bureaucracies, organizational structure and functions, excellence in organizational performance (e.g., Harrison, 1987), and interorganizational partnerships
- Societal theories—for example, theories of urban development, international behavior, cultural institutions, technological development, and marketplace functions

Other examples cut across some of these illustrative types. Decision-making theory (Carroll & Johnson, 1992), for instance, can involve individuals, organizations, or social groups. As another example, a common topic of case studies is the evaluation of publicly supported programs, such as federal, state, or local programs. In this situation, the development of a theory of how a program is supposed to work is essential to the design of the evaluation but has been commonly underemphasized in the past (Bickman, 1987). According to Bickman, analysts have frequently confused the theory of the program (e.g., how to make education more effective) with the theory of program implementation (e.g., how to install an effective program). Where policymakers want to know the desired substantive steps (e.g., describe a newly effective curriculum), the analysts unfortunately recommend managerial steps (e.g., hire a good project director). This mismatch can be avoided by giving closer attention to the substantive theory.

Generalizing from case study to theory. Theory development does not only facilitate the data collection phase of the ensuing case study. The appropriately developed theory also is the level at which the generalization of the case study results will occur. This role of theory has been characterized throughout this book as "analytic generalization" and has been contrasted with another way of generalizing results, known as "statistical generalization." Understanding the distinction between these two types of generalization may be your most important challenge in doing case studies.

Let us take the more commonly recognized way of generalizing—"statistical generalization"—first, although it is the less relevant one for doing case studies. In statistical generalization, an inference is made about a population (or universe) on the basis of empirical data collected about a sample. This is shown as a *Level One Inference* in Figure 2.2.[1] This method of generalizing

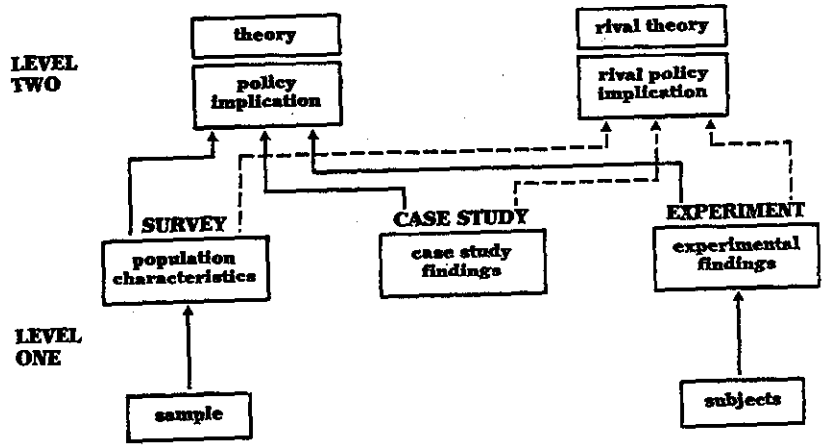

Figure 2.2. Making Inferences: Two Levels
SOURCE: COSMOS Corporation.

is commonly recognized because research investigators have ready access to formulas for determining the confidence with which generalizations can be made, depending mostly upon the size and internal variation within the universe and sample. Moreover, this is the most common way of generalizing when doing surveys (e.g., Fowler, 1988; Lavrakas, 1987), and it is an integral (though not the only) part of generalizing from experiments.

A fatal flaw in doing case studies is to conceive of statistical generalization as the method of generalizing the results of the case. This is because cases are not "sampling units" and should not be chosen for this reason. Rather, individual case studies are to be selected as a laboratory investigator selects the topic of a new experiment. Multiple cases, in this sense, should be considered like multiple experiments (or multiple surveys). Under these circumstances, the method of generalization is "analytic generalization," in which a previously developed theory is used as a template with which to compare the empirical results of the case study. If two or more cases are shown to support the same theory, replication may be claimed. The empirical results may be considered yet more potent if two or more cases support the same theory but do not support an equally plausible, *rival* theory. Graphically, this type of generalization is shown as a *Level Two Inference* in Figure 2.2.

Analytic generalization can be used whether your case study involves one or several cases, which shall be later referenced as single-case or multiple-case studies. Further, the logic of replication and the distinction between statistical

and analytic generalization will be covered in greater detail in the discussion of multiple-case study designs. The main point at this juncture is that you should try to aim toward analytic generalization in doing case studies, and you should avoid thinking in such confusing terms as "the sample of cases" or the "small sample size of cases," as if a single case study were like a single respondent in a survey or a single subject in an experiment. In other words, in terms of Figure 2.2, you should aim for *Level Two Inferences* when doing case studies.

Because of the importance of this distinction between the two ways of generalizing, you will find repeated examples and discussion throughout the remainder of this chapter as well as in Chapter 5.

Summary. This subsection has suggested that a complete research design, covering the five components described earlier, in fact requires the development of a theoretical framework for the case study that is to be conducted. Rather than resisting such a requirement, a good case study investigator should make the effort to develop this theoretical framework, no matter whether the study is to be explanatory, descriptive, or exploratory. The use of theory, in doing case studies, not only is an immense aid in defining the appropriate research design and data collection but also becomes the main vehicle for generalizing the results of the case study.

CRITERIA FOR JUDGING THE QUALITY OF RESEARCH DESIGNS

Because a research design is supposed to represent a logical set of statements, you also can judge the quality of any given design according to certain logical tests. Concepts that have been offered for these tests include trustworthiness, credibility, confirmability, and data dependability (U.S. General Accounting Office, 1990).

Four tests, however, have been commonly used to establish the quality of any empirical social research. Because case studies are one form of such empirical research, the four tests also are relevant to case study research. Therefore, an important innovation of this book is the identification of several tactics for dealing with these four tests when doing case studies. Figure 2.3 lists the four widely used tests and the recommended case study tactics as well as a cross-reference to the phase of research when the tactic is to be used. (Each tactic is later described in detail in the relevant chapter of this book.)

tests	case study tactic	phase of research in which tactic occurs
construct validity	– use multiple sources of evidence	data collection
	– establish chain of evidence	data collection
	– have key informants review draft case study report	composition
internal validity	– do pattern-matching	data analysis
	– do explanation-building	data analysis
	– do time-series analysis	data analysis
external validity	– use replication logic in multiple-case studies	research design
reliability	– use case study protocol	data collection
	– develop case study data base	data collection

Figure 2.3. Case Study Tactics for Four Design Tests
SOURCE: COSMOS Corporation.

Because the four tests are common to all social science methods, the tests have been summarized in numerous textbooks (see Kidder & Judd, 1986, pp. 26-29):

- *Construct validity*: establishing correct operational measures for the concepts being studied
- *Internal validity* (for explanatory or causal studies only, and not for descriptive or exploratory studies): establishing a causal relationship, whereby certain conditions are shown to lead to other conditions, as distinguished from spurious relationships
- *External validity*: establishing the domain to which a study's findings can be generalized
- *Reliability*: demonstrating that the operations of a study—such as the data collection procedures can be repeated, with the same results

This list is much more complex than the standard "validity" and "reliability" notions to which most students have been exposed, and each item deserves explicit attention. For case studies, an important revelation is that the several tactics to be used in dealing with these tests should be applied throughout the

subsequent conduct of the case study, and not just at the beginning. In this sense, "design work" actually continues beyond the initial design plans.

Construct Validity

This first test is especially problematic in case study research. People who have been critical of case studies often point to the fact that a case study investigator fails to develop a sufficiently operational set of measures and that "subjective" judgments are used to collect the data. Take an example such as studying "neighborhood change"—a common case study topic.

Over the years, concerns have arisen over how certain urban neighborhoods have changed their character. Any number of case studies have examined the types of changes and their consequences. However, without any prior specification of the significant, operational events that constitute "change," a reader cannot tell whether the recorded changes in a case study genuinely reflect critical events in a neighborhood or whether they happen to be based on an investigator's impressions only.

Neighborhood change can indeed cover a wide variety of phenomena: racial turnover, housing deterioration and abandonment, changes in the pattern of urban services, shifts in a neighborhood's economic institutions, or the turnover from low- to middle-income residents in "gentrifying" neighborhoods. To meet the test of construct validity, an investigator must be sure to cover two steps:

1. Select the specific types of changes that are to be studied (in relation to the original objectives of the study) and
2. Demonstrate that the selected measures of these changes do indeed reflect the specific types of change that have been selected.

For example, suppose you satisfy the first step by stating that you plan to study the rise in neighborhood crime. The second step now demands that you also justify why you might be using police-reported crime (which happens to be the standard measure used in the FBI Uniform Crime Reports) as your measure of crime. Perhaps this is not a valid measure, given that large proportions of crimes are not reported to the police.

As Figure 2.3 shows for doing case studies, three tactics are available to increase construct validity. The first is the use of *multiple sources of evidence*, in a manner encouraging convergent lines of inquiry, and this tactic is relevant during data collection (see Chapter 4). A second tactic is to establish a *chain of evidence*, also relevant during data collection (Chapter 4). The third tactic

is to have the draft case study report reviewed by key informants (a procedure described further in Chapter 6).

Internal Validity

This second test has been given the greatest attention in experimental and quasi-experimental research (see Campbell & Stanley, 1966; Cook & Campbell, 1979). Numerous "threats" to validity have been identified, mainly dealing with spurious effects. However, because so many textbooks already cover this topic, only two points need to be made here.

First, internal validity is a concern only for causal (or explanatory) case studies, in which an investigator is trying to determine whether event *x* led to event *y*. If the investigator incorrectly concludes that there is a causal relationship between x and y without knowing that some third factor—*z*—may actually have caused y, the research design has failed to deal with some threat to internal validity. Note that this logic is inapplicable to descriptive or exploratory studies (whether the studies are case studies, surveys, or experiments), which are not concerned with making causal statements.

Second, the concern over internal validity, for case study research, may be extended to the broader problem of making inferences. Basically, a case study involves an inference every time an event cannot be directly observed. Thus an investigator will "infer" that a particular event resulted from some earlier occurrence, based on interview and documentary evidence collected as part of the case study. Is the inference correct? Have all the rival explanations and possibilities been considered? Is the evidence convergent? Does it appear to be airtight? A research design that has anticipated these questions has begun to deal with the overall problem of making inferences and therefore the specific problem of internal validity.

However, the specific tactics for achieving this result are difficult to identify. This is especially true in doing case studies. As one set of suggestions, Figure 2.3 shows that the analytic tactic of *pattern-matching*, already touched upon but to be described further in Chapter 5, is one way of addressing internal validity. Two related analytic tactics, *explanation-building* and *time-series analysis*, also are described in Chapter 5.

External Validity

The third test deals with the problem of knowing whether a study's findings are generalizable beyond the immediate case study. In the simplest example,

if a study of neighborhood change focused on one neighborhood, are the results applicable to another neighborhood? The external validity problem has been a major barrier in doing case studies. Critics typically state that single cases offer a poor basis for generalizing. However, such critics are implicitly contrasting the situation to survey research, in which a "sample" (if selected correctly) readily generalizes to a larger universe. *This analogy to samples and universes is incorrect when dealing with case studies.* This is because survey research relies on *statistical* generalization, whereas case studies (as with experiments) rely on *analytical* generalization. In analytical generalization, the investigator is striving to generalize a particular set of results to some broader theory (see BOX 7).

For example, the theory of neighborhood change that led to a case study in the first place is the same theory that will help to identify the other cases to which the results are generalizable. If a study had focused on "gentrification" (see Auger, 1979), the procedure for selecting a neighborhood for study also will have identified those types of neighborhoods within which gentrification was occurring. In principle, theories about changes in all of these neighborhoods would be the target to which the results could later be generalized.

The generalization is not automatic, however. A theory must be tested through replications of the findings in a second or even a third neighborhood, where the theory has specified that the same results should occur. Once such replication has been made, the results might be accepted for a much larger number of similar neighborhoods, even though further replications have not been performed. This *replication logic* is the same that underlies the use of experiments (and allows scientists to generalize from one experiment to another) and, as shown in Figure 2.3, will be discussed further in this chapter in the section on multiple-case designs.

Reliability

Most people are probably already familiar with this final test. The objective is to be sure that, if a later investigator followed exactly the same procedures as described by an earlier investigator and conducted the same case study all over again, the later investigator should arrive at the same findings and conclusions. (Note that the emphasis is on doing the *same* case over again, not on "replicating" the results of one case by doing *another* case study.) The goal of reliability is to minimize the errors and biases in a study.

One prerequisite for allowing this other investigator to repeat an earlier case study is the need to document the procedures followed in the earlier case. Without such documentation, you could not even repeat your own work

BOX 7
How Case Studies Can Be Generalized to Theory

A common complaint about case studies is that it is difficult to generalize from one case to another. Thus analysts fall into the trap of trying to select a "representative" case or set of cases. Yet no set of cases, no matter how large, is likely to deal satisfactorily with the complaint.

The problem lies in the very notion of generalizing to other case studies. Instead, an analyst should try to generalize findings to "theory," analogous to the way a scientist generalizes from experimental results to theory. (Note that the scientist does not attempt to select "representative" experiments.)

This approach is well illustrated by Jane Jacobs in her famous book, *The Death and Life of Great American Cities* (1961). The book is based mostly on experiences from New York City. However, the chapter topics, rather than reflecting the single experiences of New York, cover broader theoretical issues in urban planning, such as the role of sidewalks, the role of neighborhood parks, the need for primary mixed uses, the need for small blocks, and the processes of slumming and unslumming. In the aggregate, these issues in fact represent the building of a theory of urban planning.

Jacob's book created heated controversy in the planning profession. As a partial result, new empirical inquiries were made in other locales to examine one or another facet of her rich and provocative ideas. Her *theory*, in essence, became the vehicle for examining other cases, and the theory still stands as a significant contribution to the field of urban planning.

(which is another way of dealing with reliability). In the past, case study research procedures have been poorly documented, making external reviewers suspicious of the reliability of the case study. As remedies, Chapter 3 will discuss the use of a *case study protocol* to deal with the documentation problem in detail, and Chapter 4 will describe another tactic, the development of a *case study database*.

The general way of approaching the reliability problem is to make as many steps as operational as possible and to conduct research as if someone were always looking over your shoulder. In accounting and bookkeeping, one is always aware that any calculations must be capable of being audited. In this sense, an auditor is also performing a reliability check and must be able to produce the same results if the same procedures are followed. A good guideline for doing case studies is therefore to conduct the research so that an auditor could repeat the procedures and arrive at the same results.

Summary. Four tests may be considered relevant in judging the quality of a research design. In designing and doing case studies, various tactics are available to deal with these tests, though not all of the tactics occur at the formal stage of designing a case study. Some of the tactics occur during the data collection, data analysis, or compositional phases of the research and are therefore described in greater detail in subsequent chapters of this book.

CASE STUDY DESIGNS

These general characteristics of research designs serve as a background for considering the specific designs for case studies. Four types of designs will be discussed, based on a 2 × 2 matrix (see Figure 2.4). The matrix assumes that single- and multiple-case studies reflect different design situations and that, within these two types, there also can be a unitary or multiple units of analysis. Thus, for the case study strategy, the four types of designs are (a) single-case (holistic) designs, (b) single-case (embedded) designs, (c) multiple-case (holistic) designs, and (d) multiple-case (embedded) designs. The rationale for these four types of designs is as follows.

What Are the Potential Single-Case Designs?

Rationale for single-case designs. A primary distinction in designing case studies is between *single-* and *multiple*-case designs. This means the need for a decision, prior to any data collection, on whether a single-case study or multiple cases are going to be used to address the research questions.

The single-case study is an appropriate design under several circumstances. First, recall that a single-case study is analogous to a single experiment, and many of the same conditions that justify a single experiment also justify a single-case study. One rationale for a single case is when it represents the *critical case* in testing a well-formulated theory (again, note the analogy to the critical experiment). The theory has specified a clear set of propositions as well as the circumstances within which the propositions are believed to be true. To confirm, challenge, or extend the theory, there may exist a single case, meeting all of the conditions for testing the theory. The single case can then be used to determine whether a theory's propositions are correct or whether some alternative set of explanations might be more relevant. In this manner, like Graham Allison's comparison of three theories of bureaucratic functioning and the Cuban missile crisis (described in Chapter 1, BOX 2), the single case can represent a significant contribution to knowledge and

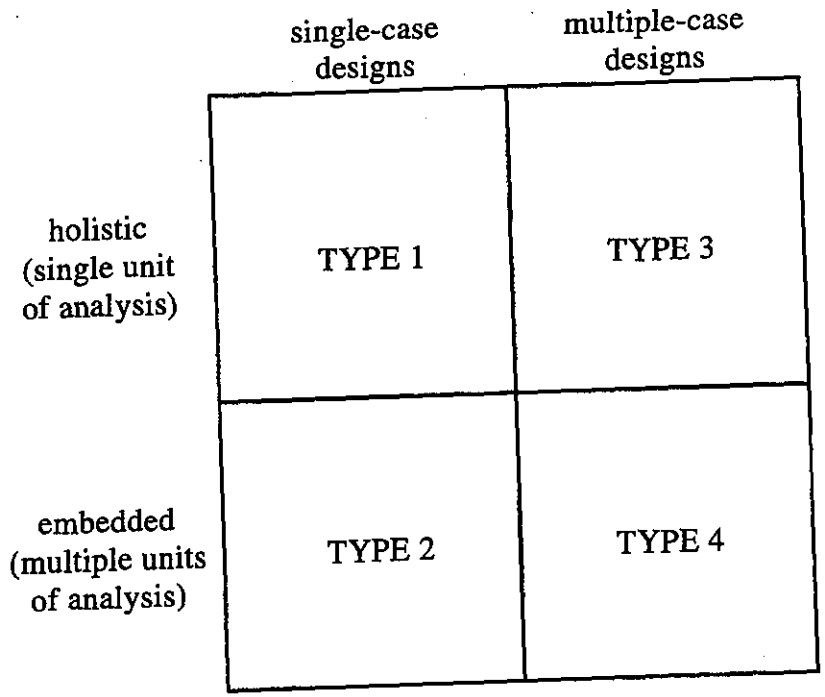

Figure 2.4. Basic Types of Designs for Case Studies
SOURCE: COSMOS Corporation.

theory-building. Such a study can even help to refocus future investigations in an entire field. (See BOX 8 for another example, in the field of organizational innovation.)

A second rationale for a single case is one in which the case represents an *extreme or unique case*. This has commonly been the situation in clinical psychology, in which a specific injury or disorder may be so rare that any single case is worth documenting and analyzing. For instance, one rare clinical syndrome is the inability of certain clinical patients to recognize familiar faces. Given visual cues alone, such patients are unable to recognize loved ones, friends, pictures of famous people, or (in some cases) their own image in a mirror. This syndrome appears to be due to some physical injury to the brain. Yet the syndrome occurs so rarely that scientists have been unable to establish any common patterns (Yin, 1970, 1978). In such circumstances, the single-case study is an appropriate research design whenever a new person with this syndrome—known as prosopagnosia—is encountered. The case

BOX 8
The Single Case Study as the Critical Case

One rationale for selecting a single-case rather than a multiple-case design is that the single case represents the *critical test of a significant theory*. Neal Gross et al. used such a design by focusing on a single school in their book, *Implementing Organizational Innovations* (1971).

The school was selected because it had a prior history of innovation and could not be claimed to suffer from "barriers to innovation." In the prevailing theories, such barriers had been prominently cited as the major reason that innovations failed. Gross et al. showed that, in this school, an innovation also failed but that the failure could not be attributed to any barriers. Implementation processes, rather than barriers, appeared to account for the outcomes.

In this manner, the book, though limited to a single case, represents a watershed in innovation theory. Prior to the study, analysts had focused on the identification of barriers; since the study, the literature has been much more dominated by studies of the implementation process.

study would document the person's abilities and disabilities to determine the precise nature of the face recognition deficit but also to ascertain whether related disorders exist.

A third rationale for a single case study is the *revelatory case*. This situation exists when an investigator has an opportunity to observe and analyze a phenomenon previously inaccessible to scientific investigation, such as Whyte's *Street Corner Society*, previously described in Chapter 1, BOX 1. A latter-day example is Elliot Liebow's famous case study of unemployed blacks, *Tally's Corner* (see BOX 9). Liebow had the opportunity to meet the men in one neighborhood in Washington, DC, and to learn about their everyday lives. His observations of and insights into the problems of unemployment formed a significant case study, because few social scientists had previously had the opportunity to investigate these problems, even though the problems were common across the country (as distinguished from the rare or unique case). When other investigators have similar types of opportunities and can uncover some prevalent phenomenon previously inaccessible to scientists, such conditions justify the use of a single-case study on the grounds of its revelatory nature.

These three rationales serve as the major reasons for conducting a single-case study. There are other situations in which the single-case study may be conducted as a prelude to further study, such as the use of case studies as

BOX 9
The Revelatory Case as a Single Case

Another rationale for selecting a single-case rather than a multiple-case design is that the investigator has access to a situation previously inaccessible to scientific observation. The case study is therefore worth conducting because the descriptive information alone will be revelatory.

Such was the situation in Elliot Liebow's sociological classic, *Tally's Corner* (1967). The book is about a single group of black men, living in a poor, inner-city neighborhood. By befriending these men, the author was able to learn about their lifestyles, their coping behavior, and in particular their sensitivity to unemployment and failure. The book provides insights into a subculture that has prevailed in many U.S. cities for a long period of time, but one that had been only obscurely understood. The single case showed how investigations of such topics could be done, stimulating much further research and eventually the development of policy actions.

exploratory devices or such as the conduct of a pilot case that is the first of a multiple-case study. However, in these latter instances, the single-case study cannot be regarded as a complete study on its own.

Whatever the rationale for doing single cases (and there may be more than the three mentioned here), a potential vulnerability of the single-case design is that a case may later turn out not to be the case it was thought to be at the outset. Single-case designs therefore require careful investigation of the potential case to minimize the chances of misrepresentation and to maximize the access needed to collect the case study evidence. A fair warning is not to commit oneself to the single case until these major concerns have been covered.

Holistic versus embedded case studies. The same case study may involve *more than one unit of analysis*. This occurs when, within a single case, attention also is given to a subunit or subunits (see BOX 10). For instance, even though a case study might be about a single public program, the analysis might include outcomes from individual projects within the program (and possibly even some quantitative analyses of large numbers of projects). In an organizational study, the embedded units also might be "process" units—such as meetings, roles, or locations. In either situation, these embedded units can be selected through sampling or cluster techniques (McClintock, 1985). However the units are selected, the resulting design

BOX 10
An Embedded, Single-Case Design

Union Democracy (1956) is a highly regarded case study by three eminent academicians—Seymour Martin Lipset, Martin Trow, and James Coleman. The case study is about the inside politics of the International Typographical Union and involves *several* units of analysis (see the following table). The main unit was the organization as a whole, the smallest unit was the individual member, and several intermediary units also were important. At each level of analysis, different data collection techniques were used, ranging from historical to survey analysis.

would be called an *embedded case study design* (see Figure 2.4, Type 2). In contrast, if the case study examined only the global nature of a program or of an organization, a *holistic design* would have been used (see Figure 2.4, Type 1).

Both variations of single-case studies have different strengths and weaknesses. The holistic design is advantageous when no logical subunits can be identified and when the relevant theory underlying the case study is itself of a holistic nature. Potential problems arise, however, when a global approach allows an investigator to avoid examining any specific phenomenon in operational detail. Another typical problem with the holistic design is that the entire case study may be conducted at an abstract level, lacking any clear measures or data.

A further problem with the holistic design is that the entire nature of the case study may shift, unbeknownst to the researcher, during the course of study. The initial study questions may have reflected one orientation, but as the case study proceeds, a different orientation may emerge, and the evidence begins to address different questions. Although some people have claimed such flexibility to be a strength of the case study approach, in fact, the largest criticism of case studies is based on this type of shift—in which the original research design is no longer appropriate for the research questions being asked (see Yin, Bateman, & Moore, 1983). Because of this problem, such unsuspected slippage needs to be avoided; if the relevant research questions really do change, you should simply start over again, with a new research design. One way to increase the sensitivity to such slippage is to have a set of subunits. With such subunits, an embedded design can serve as an important device for focusing a case study inquiry.

Kinds of Data (BOX 10 Continued)

Unit Being Characterized	*Total System*	*Intermediate Units*		*Individuals*	
	Issues, Data on Occupation; Union Laws; Policies; Historical Data; Convention Reports	*Locals' Histories and Voting Records: Issues on Local Level; Size of Locals*	*Shops' Voting Records: Shop Size*	*Interviews with Leaders*	*Interviews of the Sample of Men*
ITU as a whole	Structural, environmental, behavioral properties	By inference, communication network (structural)			
Locals	Behavioral properties (militancy, etc.)	Behavioral properties, size	By inference, communication network (structural)	Structural, environmental, behavioral properties	
Shops			Behavioral properties, size		Distributions of individual properties
Other immediate social environment of men	The social climate, by inference from dominant issues and election outcome	The social climate, by inference from dominant issues and election outcome			Chapel chairman's attributes; friends' attributes
Men	By inference, dominant values and interests	By inference: values, interests, and loyalties (e.g., local over international)	By inference: values, interests, loyalties (e.g., to shop over local)	By inference: values	Behavior, background, values, attitudes

SOURCE: Lipset, Trow, & Coleman (1956, p. 422). Reprinted by permission.

An embedded design, however, also has some pitfalls. A major one occurs when the case study focuses only on the subunit level and fails to return to the larger unit of analysis. A program evaluation that includes project characteristics as a subunit of analysis, for instance, becomes a project study if no investigating is done at the larger unit—that is, the "program." Similarly, a study of organizational climate may involve individual employees as a subunit of study. However, if the data focus only on individual employees, the study will in fact become an employment and not an organizational study. What has happened is that the original phenomenon of interest (organizational climate) has become the context and not the target of study.

Summary. Single cases are a common design for doing case studies, and two types have been described: those using holistic designs and those using embedded units of analysis. Overall, the single-case design is eminently justifiable under certain conditions—where the case represents a critical test of existing theory, where the case is a rare or unique event, or where the case serves a revelatory purpose.

A major step in designing and conducting a single case is defining the unit of analysis (or the case itself). An operational definition is needed and some precaution must be taken—before a total commitment to the whole case study is made—to ensure that the case in fact is relevant to the issues and questions of interest.

Within the single case may still be incorporated subunits of analyses, so that a more complex—or embedded—design is developed. The subunits can often add significant opportunities for extensive analysis, enhancing the insights into the single case. However, if too much attention is given to these subunits, and if the larger, holistic aspects of the case begin to be ignored, the case study itself will have shifted its orientation and changed its nature. This shift might in fact be justifiable, but it should not come as a surprise to the investigator.

What Are the Potential Multiple-Case Designs?

The same study may contain more than a single case. When this occurs, the study has to use a multiple-case design, and such designs have increased in frequency in recent years. A common example is a study of school innovations (such as open classrooms, teacher aides, or new technology) in which independent innovations occur at different sites. Thus each site might be the subject of an individual case study, and the study as a whole would have used a multiple-case design.

Multiple- versus single-case designs. In some fields, multiple-case studies have been considered a different "methodology" than single-case studies. For example, both anthropology and political science have developed one set of rationales for doing single-case studies and a second set for doing what have been considered "comparative" (or multiple-case) studies (see Eckstein, 1975; George, 1979). From the perspective of this book, however, the choice between single- and multiple-case designs remains within the same methodological framework—and no broad distinction is made between the so-called classic (that is, single-) case study and multiple-case studies. The choice is considered one of research design, with both being included under the case study strategy.

Multiple-case designs have distinct advantages and disadvantages in comparison with single-case designs. The evidence from multiple cases is often considered more compelling, and the overall study is therefore regarded as being more robust (Herriott & Firestone, 1983). At the same time, the rationale for single-case designs usually cannot be satisfied by multiple cases. The unusual or rare case, the critical case, and the revelatory case are all likely to involve only single cases, by definition. Moreover, the conduct of a multiple-case study can require extensive resources and time beyond the means of a single student or independent research investigator.

Therefore, the decision to undertake multiple-case studies cannot be taken lightly. Every case should serve a specific purpose within the overall scope of inquiry. Here, *a major insight is to consider multiple cases as one would consider multiple experiments*—that is, to follow a "replication" logic. This is far different from a mistaken analogy in the past, which incorrectly considered multiple cases to be similar to the multiple respondents in a survey (or to the multiple subjects *within* an experiment)—that is, to follow a "sampling" logic. The methodological differences between these two views are revealed by the different rationales underlying the replication as opposed to sampling logics.

Replication, not sampling logic, for multiple-case studies. The replication logic is analogous to that used in multiple experiments (see Hersen & Barlow, 1976). Thus, if one has access only to three cases of a rare, clinical syndrome in psychology or medical science, the appropriate research design is one in which the same results are predicted for each of the three cases, thereby producing evidence that the three cases did indeed involve the same syndrome. If similar results are obtained from all three cases, replication is said to have taken place. This replication logic is the same whether one is repeating certain critical experiments, is limited to a few cases due to the expense or difficulty in performing a surgical preparation in animals, or is

limited by the rarity of occurrence of a clinical syndrome. In each of these situations, an individual case or subject is considered akin to a single experiment, and the analysis must follow cross-experiment rather than *within*-experiment design and logic.

The logic underlying the use of multiple-case studies is the same. Each case must be carefully selected so that it either (a) predicts similar results (a *literal replication*) or (b) produces contrasting results but for predictable reasons (a *theoretical replication*). The ability to conduct six or ten case studies, arranged effectively within a multiple-case design, is analogous to the ability to conduct six to ten experiments on related topics; a few cases (two or three) would be literal replications, whereas a few other cases (four to six) might be designed to pursue two different patterns of theoretical replications. If all the cases turn out as predicted, these six to ten cases, in the aggregate, would have provided compelling support for the initial set of propositions. If the cases are in some way contradictory, the initial propositions must be revised and retested with another set of cases. Again, this logic is similar to the way scientists deal with contradictory experimental findings.

An important step in all of these replication procedures is the development of a rich, theoretical framework. The framework needs to state the conditions under which a particular phenomenon is likely to be found (a literal replication) as well as the conditions when it is not likely to be found (a theoretical replication). The theoretical framework later becomes the vehicle for generalizing to new cases, again similar to the role played in cross-experiment designs. Furthermore, just as with experimental science, if some of the empirical cases do not work as predicted, modification must be made to the theory. Remember, too, that theories can be practical, and not just academic. The study in BOX 11 contains an excellent example of a multiple-case study (two cases) whose cases and conclusions are tied together by a practical, policy-oriented theory.

To take another example, one might consider the initial proposition that an increase in microcomputer use in school districts will occur when such technologies are used for both administrative and instructional applications, but not either alone. To pursue this proposition in a multiple-case study design, three or four cases might be selected in which both types of applications are present, to determine whether, in fact, microcomputer use did increase over a period of time (the investigation would be predicting a literal replication in these three or four cases). Three or four additional cases might be selected in which only administrative applications are present, with the prediction being little increase in use (predicting a theoretical replication). Finally, three or four other cases would be selected in which only instructional applications

BOX 11
Multiple-Case Studies and a Policy-Oriented Theory

The international marketplace of the 1970s and 1980s was marked by Japan's prominence. Much of its strength was attributable to the role of centralized planning and support by government agencies. In contrast, the United States was considered to have no counterpart support structures. Gregory Hooks's excellent case study (1990) points to a counterexample, frequently ignored by advocates: the role of the U.S. defense department in implementing an industrial planning policy within defense-related industries.

Hooks provides quantitative data on two cases—the aeronautics industry and the microelectronics industry. One industry was much more dependent upon government than the other. However, in both cases, Hooks's evidence shows how the defense department supported the development of these industries through financial support, ensuring demand, and support of R&D.

are present, with the same prediction of little increase in use, but for different reasons than the administrative-only cases (another theoretical replication). If this entire pattern of results across these multiple cases is indeed found, the nine to twelve cases, in the aggregate, would provide substantial support for the initial proposition. (See BOX 12 for another example of a multiple-case replication design, but from the field of urban studies.)

This replication logic, whether applied to experiments or to case studies, must be distinguished from the sampling logic commonly used in surveys. According to the sampling logic, a number of respondents (or subjects) are assumed to "represent" a larger pool of respondents (or subjects), so that data from a smaller number of persons are assumed to represent the data that might have been collected from the entire pool.

The sampling logic demands an operational enumeration of the entire universe or pool of potential respondents and then a statistical procedure for selecting the specific subset of respondents to be surveyed. This logic is applicable whenever an investigator is interested in determining the prevalence or frequency of a particular phenomenon and when it is too expensive or impractical to survey the entire universe or pool. The resulting data from the sample that is actually surveyed are assumed to reflect the entire universe or pool, with inferential statistics used to establish the confidence intervals for which this representation is actually accurate.

BOX 12
A Multiple-Case, Replication Design

A common problem of the 1960s and 1970s was how to get good advice to city governments. Peter Szanton's book, *Not Well Advised* (1981), reviewed the experiences of numerous attempts by university and research groups to collaborate with city officials.

The book is an excellent example of a multiple-case, replication design. Szanton starts with eight case studies, showing how different university groups all failed to help cities. The eight cases are sufficient "replications" to convince the reader of a general phenomenon. Szanton then provides five more case studies, in which nonuniversity groups also failed, concluding that failure was therefore not necessarily inherent in the academic enterprise. Yet a third group of cases shows how university groups have successfully helped business, engineering firms, and sectors other than city government. A final set of three cases shows that those few groups able to help city government were concerned with implementation and not just with the production of new ideas, leading to the major conclusion that city governments may have peculiar needs in receiving advice.

Within each of the four groups of case studies, Szanton has illustrated the principle of *literal* replication. Across the four groups, he has illustrated *theoretical* replication. This potent case study design can and should be applied to many other topics.

Any application of this sampling logic to case studies would be misplaced. First, case studies should not generally be used to assess the incidence of phenomena. Second, a case study would have to cover both the phenomenon of interest and its context, yielding a large number of potentially relevant variables. In turn, this would require an impossibly large number of cases—too large to allow any statistical consideration of the relevant variables.

Third, if a sampling logic had to be applied to all types of research, many important topics could not be empirically investigated, such as in the following problem: Your investigation deals with the role of the presidency of the United States, and you are interested in studying the behavior of the incumbent from some leadership perspective. The leadership perspective, to be at all faithful to the complexity of reality, must incorporate dozens if not hundreds of relevant variables. Any sampling logic simply would be misplaced under such circumstances, as there have been only 42 presidents since the beginning of the Republic. Moreover, you would probably not have the resources to

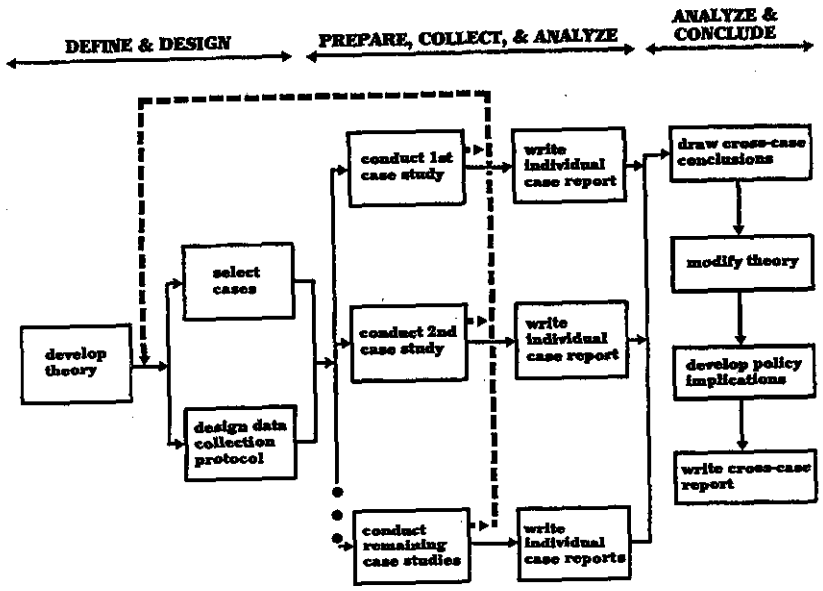

Figure 2.5. Case Study Method
SOURCE: COSMOS Corporation.

conduct a full study of all 42 incumbents (and even if you did, you would still have too many variables in relation to the 42 data points available). This type of study just could not be done, following the sampling logic; if the replication logic is followed, however, the study is eminently feasible.

The replication approach to multiple-case studies is illustrated in Figure 2.5. (This figure is derived from research on the case study method; see Yin, Bateman, & Moore, 1983.) The figure indicates that the initial step in designing the study must consist of theory development and then shows that case selection and the definition of specific measures are important steps in the design and data collection process. Each individual case study consists of a "whole" study, in which convergent evidence is sought regarding the facts and conclusions for the case; each case's conclusions are then considered to be the information needing replication by other individual cases. Both the individual cases and the multiple-case results can and should be the focus of a summary report. For each individual case, the report should indicate how and why a particular proposition was demonstrated (or not demonstrated). Across cases, the report should indicate the extent of the replication logic and

why certain cases were predicted to have certain results, whereas other cases—if any—were predicted to have contrasting results.

Again, Figure 2.5 depicts a very different logic than that of sampling design. This is a difficult step to perceive and is worth extensive discussion with colleagues before proceeding with any case study design.

When using a multiple-case design, a further question you will encounter has to do with the *number* of cases deemed necessary or sufficient for your study. However, because a sampling logic should not be used, the typical criteria regarding sample size also are irrelevant. Instead, you should think of this decision as a reflection of the number of case replications—both literal and theoretical—that you would like to have in your study.

For the number of literal replications, an appropriate analogy from statistical studies is the selection of the criterion for establishing levels of significance. Much as the choice of "$p < .05$" or "$p < .01$" is not derived from any formula but is a matter of discretionary, judgmental choice, the selection of the number of replications depends upon the certainty you want to have about your multiple-case results (as with the higher criterion for establishing statistical significance, the greater certainty lies with the larger number of cases). For example, you may want to settle for two or three literal replications when the rival theories are grossly different and the issue at hand does not demand an excessive degree of certainty. However, if your rivals have subtle differences or if you want a high degree of certainty, you may press for five, six, or more replications.

For the number of theoretical replications, the important consideration is related to your sense of the complexity of the realm of external validity. When you are uncertain whether external conditions will produce different case study results, you may want to articulate these relevant conditions more explicitly at the outset of your study and identify a larger number of cases to be included. For example, in the neighborhood example used previously in discussing external validity (see the section titled "External Validity"), a common concern from the standpoint of policy research (e.g., Majchrzak, 1984) is that ethnically and racially different neighborhoods do not usually follow similar courses of change. A study of gentrification would therefore want to include at least some number of cases that varied along ethnic or racial lines (and *within* each type of case, one would still want a minimum of two or three literal replications). In contrast, when external conditions are not thought to produce much variation in the phenomenon being studied, a smaller number of theoretical replications is needed.

Multiple-case studies: Holistic or embedded. The fact that a design calls for multiple-case studies does not eliminate the variation identified earlier with single cases: Each individual case may still be holistic or embedded. In other words, a multiple-case study may consist of multiple holistic cases (see Figure 2.4, Type 3) or of multiple embedded cases (see Figure 2.4, Type 4).

The difference between these two designs depends upon the type of phenomenon being studied. In an embedded design, a study may even call for the conduct of a survey at each case study site. For instance, supposing a study is concerned with the delivery of services by different community mental health centers (facilities) (see Larsen, 1982). Each center may rightfully be the topic of a case study; the theoretical framework may dictate that nine such centers be included as case studies, three to replicate a direct result (literal replication) and six others to deal with contrasting conditions (theoretical replications).

In all nine centers, an embedded design is used because surveys of the centers' clients will be conducted. However, the results of each survey will *not* be pooled across centers. Rather, the survey data will be part of the findings for each individual center, or case. These data may be highly quantitative, focusing on the attitudes and behavior of individual clients, and the data will be used along with archival information to interpret the success and operations at the given center. If, in contrast, the survey data are pooled across centers, a multiple-case study design is no longer being used, and the investigation is likely to be using a survey rather than case study design.

Summary. This section has dealt with situations in which the same investigation may call for multiple-case studies. These types of designs are becoming more prevalent, but they are more expensive and time-consuming to conduct.

Any use of multiple-case designs should follow a replication, not a sampling, logic, and an investigator must choose each case carefully. The cases should serve in a manner similar to multiple experiments, with similar results (a literal replication) or contrasting results (a theoretical replication) predicted explicitly at the outset of the investigation.

The replication design does not necessarily mean that each case study needs to be either holistic or embedded. The individual cases, within a multiple-case study design, may be either. When an embedded design is used, each individual case study may in fact include the collection and analysis of highly quantitative data, including the use of surveys within each case.

How Case Study Designs Can Be Kept Flexible

A final reminder is that a case study design is not something completed only at the outset of a study. The design can be altered and revised after the initial stages of a study, but only under stringent circumstances.

As an example, *pilot* case studies may reveal inadequacies in the initial design or may help to articulate it. In the event of a single-case design, what was thought to be a revelatory or unique case may not turn out to be so after all. In the event of a multiple-case design, the selection of cases may have to be modified because of new information about the cases. In other words, after some early data collection and analysis, an investigator has every right to conclude that the initial design was faulty and to modify the design. This is an appropriate and desirable use of pilot studies. (Also see Chapter 3 for more on pilot case studies.)

At the same time, an investigator must be careful not to shift, unknowingly, the theoretical concerns or objectives. If these, rather than the cases themselves, are changed, the investigator can correctly be accused of exercising a bias in conducting the research and interpreting the findings. The point is that the flexibility of case study designs is *in selecting cases different from those initially identified* (with appropriate documentation of this shift) but not in changing the purpose or objectives of the study to suit the case(s) that were found. The former situation is much like changing experiments when it is obvious that an experimental procedure is infeasible; the latter is a more subtle but still illegitimate change.

EXERCISES

1. *Defining the boundaries of a case study.* Select a topic for a case study you would like to do. Identify some basic questions to be answered by your case study. Does the naming of these questions clarify the boundaries of your case, with regard to the relevant length of time for which evidence is to be collected? The relevant organization or geographic area? The type of evidence that should be collected? The priorities for doing analysis?

2. *Defining the unit of analysis for a case study.* Examine or read the case study *The Soul of a New Machine.* What is the main unit of analysis in this book? What alternatives did you consider, and why did you select the unit that you did? Carry out the same exercise for some other case study of your choosing.

3. *Defining a case study research design.* Select one of the case studies described in the BOXES of this book. Describe the research design of this case study. How did it justify the relevant evidence to be sought, given the basic research questions to be answered? What methods were used to draw conclusions, based on the evidence? Is the design a single- or multiple-case design? Is it holistic or does it have embedded units of analysis?

4. *Establishing the rationale for single- and multiple-case studies.* Name the rationales for using a single-case study design, then name those for using a multiple-case design. Give examples of each type of design, either from the case studies described in the BOXES of this book or from other case studies with which you are acquainted. What are the advantages of each type of design?

5. *Defining the criteria for judging the quality of research designs.* Define the four criteria for judging the quality of research designs: (a) construct validity, (b) internal validity, (c) external validity, and (d) reliability. Give an example of each type of criterion in a case study you might want to do.

NOTE

1. Figure 2.2 focuses only on the formal research design process, not on data collection activities. For all three types of research, data collection techniques might be depicted as a third level and also can involve inferences—for example, for case studies this might include searching for patterns among converging types of evidence, as described further in Chapter 5; similar data collection techniques can be described for surveys or experiments—for example, questionnaire design for surveys or stimulus presentation strategies for experiments.