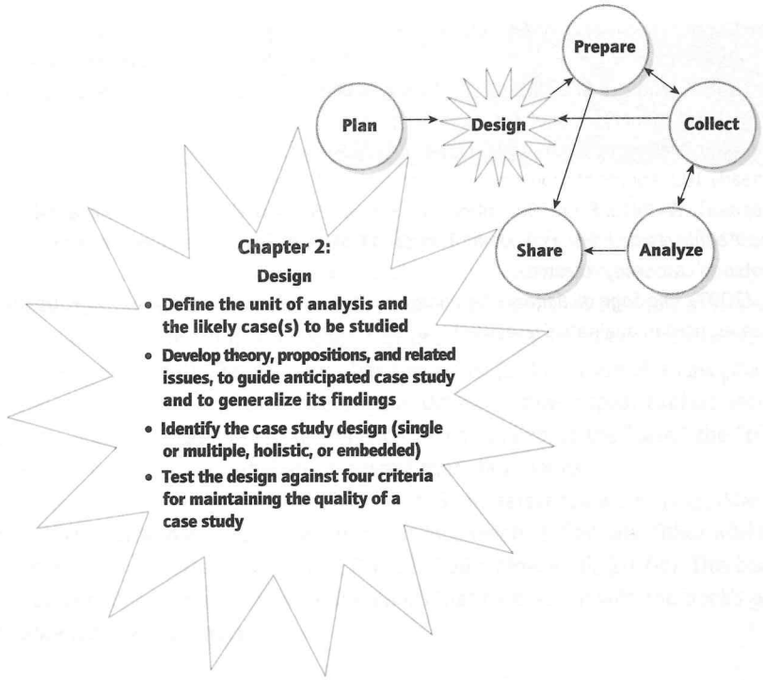


ABSTRACT

A research design is the logic that links the data to be collected (and the conclusions to be drawn) to the initial questions of study. Every empirical study has an implicit, if not explicit, research design. Articulating a "theory" about what is being studied and what is to be learned helps to strengthen a research design when doing case study research. Good theoretical propositions also lay the groundwork for generalizing the findings from the case study to other situations, by making *analytic* rather than *statistical generalizations*.

Critical to the design will be to define the "case" or unit of analysis to be studied, as well as to set some limits or bounds to the case. You can then examine the quality of your emerging design in relation to four tests commonly used in social science research: (a) construct validity, (b) internal validity, (c) external validity, and (d) reliability.

Among the specific case study designs, four major types follow a 2 × 2 matrix. The first pair consists of single-case and multiple-case designs. The second pair, which can occur in combination with either of the first pair, distinguishes between holistic and embedded designs. Whether holistic or embedded, in a multiple-case study, the selection of the cases should follow a replication rather than sampling logic. Although single-case studies can yield invaluable insights, most multiple-case study designs are likely to be stronger than single-case study designs. Trying to use even a "two-case" design is therefore a worthy objective, compared to doing a single-case study. Case study research also can be used in combination with other methods, as part of a larger mixed methods study.

2

DESIGNING CASE STUDIES

Identifying Your Case(s) and Establishing the Logic of Your Case Study

Chapter 1 has shown when you might choose to do case study research, as opposed to other types of research, to carry out a new study. The next step is to design your case study. For this purpose, as in designing any other type of research investigation, you need a plan or *research design*.

The development of this research design calls for careful craftwork. Unlike other research methods, a comprehensive and standard catalog of research designs for case study research has yet to emerge. 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. Nor have any common case study designs emerged—such as the *panel studies,* for example—used in 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 methods, such as quasi-experiments (e.g., Campbell & Stanley, 1966; Cook & Campbell, 1979). For a long time, scholars incorrectly thought that the case study was but one type of quasi-experimental design (the

"one-shot post-test-only" design—Campbell & Stanley, 1966, pp. 6–7). Although the misperception lingers to this very day, it was later corrected when one of the original authors made the following statement in the revision to the original work 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, post-test-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, case study research is a separate method that has its own research designs.

Tip: *How should I select the case(s) for my case study?*

You need sufficient access to the data for your potential case–whether to interview people, review documents or records, or make field observations. Given such access to more than a single candidate case, you should choose the case(s) that will most likely illuminate your research questions. Absent such access, you should consider changing your research questions, hopefully leading to new candidates to which you do have access.

Do you think access should be so important?

Unfortunately, case study research designs have not been codified. The following chapter therefore expands on the methodological ground broken by earlier editions 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 study 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 *a logical 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. (Nachmias & Nachmias, 1992, pp. 77–78, emphasis added)

Another way of thinking about a research design is as a "blueprint" for your 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 (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 properly 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 an organization's relationships from the vantage point of only one organization, you cannot draw unbiased conclusions about the relationships. 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

In case study research, five components of a research design are especially important:

1. a case 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, which 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 method to be used. Case study research 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.

More troublesome may be your having to come up with the substance of the questions. Many students take an initial stab, only to be discouraged when they find the same question(s) already well covered by previous research. Other less desirable questions focus on too trivial or minor parts of an issue. A helpful hint is to move in three stages. In the first, try to use the literature to narrow your interest to a key topic or two, not worrying about any specific research questions. In the second, examine closely—even dissect—a few key studies on your topic of interest. Identify the questions in those few studies and whether they conclude with new questions or loose ends for future research.

These may then stimulate your own thinking and imagination, and you may find yourself articulating some potential questions of your own. In the third stage, examine another set of studies on the same topic. They may reinforce the relevance and importance of your potential questions or even suggest ways of sharpening them.

Study propositions. As for the second component, each proposition directs attention to something that should be examined within the scope of study. For instance, assume that your research, on the topic of interorganizational partnerships, began with the following question: How and why do organizations collaborate with one another to provide joint services (for example, a manufacturer and a retail outlet collaborating to sell certain computer products)? These "how" and "why" questions, capturing what you are really interested in addressing, led you to case study research as the appropriate method in the first place. Nevertheless, these "how" and "why" questions do not sufficiently 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, besides 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 (that is, 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 methods alike—in which a topic is the subject of "exploration." Every exploration, however, should still have some purpose. Instead of propositions, the design for an exploratory study should state this purpose, as well as the criteria by which an exploration will be judged successful (or not). Consider the analogy in BOX 5 for exploratory case studies. Can you imagine how you would ask for support from Queen Isabella to do your exploratory study?

BOX 5
"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 it. 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.

Unit of analysis—the "case." This third component is related to the fundamental problem of defining the "case" to be studied—a problem that rightfully confronts many researchers at the outset of their case studies (e.g., Ragin & Becker, 1992). You will need to consider at least two different steps: defining the case and bounding the case.

In *defining the case,* for instance, the classic case studies usually focus on an individual person as the case (e.g., Bromley, 1986, p. 1). Jennifer Platt (1992) has noted how the early case studies by scholars in the Chicago school of sociology were life histories of such persons 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 the relevant individual would be collected, and several such individuals or "cases" might be included in a multiple-case study.

You would still need study questions and study propositions to help identify the relevant information to be collected about this individual or individuals. Without such questions and propositions, you might be tempted to cover "everything" about the individual(s), which is impossible to do. For example, the propositions in studying these individuals might be limited to the influence of early childhood or the role of peer relationships. Such seemingly general topics nevertheless represent a vast narrowing of the relevant data. The more a case study contains specific questions and propositions, the more it will stay within feasible limits.

Of course, the "case" also can be some event or entity other than a single individual. Case studies have been done about a broad variety of topics, including small groups, communities, decisions, programs, organizational change, and specific events. Feagin et al. (1991) contains some classic examples of these single cases in sociology and political science.

Beware of these types of cases—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 preexisted 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. Similarly, you might at first identify a specific locale, such as a "city," as your case. However, your research questions and data collection might in fact be limited to tourism in the city, city policies, or city government. These choices would differ from defining the geographic city and its population as your case.

As a general guide, the tentative definition of your case (or of the unit of analysis) is related to the way you define your initial research question(s). Suppose, for example, you want to study the role of the United States in the global economy. Years ago, Peter Drucker (1986) wrote a provocative essay (but not a case study) about fundamental changes in the world economy, including the importance of "capital movements" independent of the flow of goods and services. If you were interested in doing a case study on this topic, Drucker's work would only serve as a starting point. You would still need

to define the research question(s) of interest to you, and each question might point to a different unit of analysis (or "case"). Depending upon your question(s), the appropriate case might be a country's economy, an industry in the world marketplace, an economic policy, or the trade or capital flow between countries. Each unit of analysis and its related questions and propositions would call for a different case study, each having its own research design and data collection strategy.

If your research questions do not lead to the favoring of one unit of analysis over another, your questions may be too vague or too numerous—and you may have trouble doing a case study. However, when you do eventually arrive at a definition of the unit of analysis, do not consider closure permanent. Your choice of the unit of analysis, as with other facets of your research design, can be revisited as a result of discoveries during your data collection (see discussion and cautions about maintaining an adaptive posture, throughout this book and at the end of this chapter).

Sometimes, the unit of analysis may have been defined one way, even though the phenomenon being studied actually follows a different definition. Most frequently, investigators have confused case studies of neighborhoods with case studies of small groups (as another example, confusing a new technology with the workings of an engineering team in an organization; see BOX 6A). How a geographic *area* such as a neighborhood copes with racial transition, upgrading, and other phenomena can be quite different from how a small *group* copes with these same phenomena. For instance, *Street Corner Society* (Whyte, 1943/1993; see BOX 2A in Chapter 1 of this book) and *Tally's Corner* (Liebow, 1967; see BOX 9, this chapter) often have 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 the small groups lived in a small area with clear neighborhood definitions if not boundaries). In contrast, BOX 6B presents a good example of how units of analyses can be defined in a more discriminating manner—in the field of world trade.

BOX 6
Defining the Unit of Analysis

6A. 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 with one produced by a direct competitor, Digital Equipment Corporation (also see BOX 29, Chapter 5, p. 144).

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

(Continued)

(Continued)

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" being studied the minicomputer, or is it about the dynamics of a small group–the engineering team? The answer is critical for understanding how the case study might relate to any 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.

6B. A Clearer Choice among Units of Analysis

Ira Magaziner and Mark Patinkin's (1989) book, *The Silent War: Inside the Global Business Battles Shaping America's Future,* presents nine individual case studies (also see BOX 36, Chapter 5, p. 165). Each case 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 covers a firm–the Korean firm Samsung–and the critical policies that make it 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 covers a country–Singapore–and the policies that make it competitive. Within the country case study is an embedded unit of analysis–the development of an Apple computer factory in Singapore, serving as an illustrative example of how the national policies affect foreign investments.

To reduce the confusion and ambiguity in defining the unit of analysis or "case," one recommended practice is to discuss your potential case selection with a colleague. Try to explain to that person what questions you are trying to address and why you have chosen a specific case or group of cases as a way of addressing 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—sometimes called *bounding the case*—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 of it (the context for the case study). Similarly, if the case is about the local services in a specific geographic area, you need to decide which services to cover. Also desirable, for almost any topic that might be chosen, are the specific time boundaries to define the estimated beginning and ending of the case, for the purposes of your study (e.g., whether to include the entire or only some part of the life cycle of the entity that is to be the case).

Bounding the case in these ways will help to determine the scope of your data collection and, in particular, how you will distinguish data about the subject of your case study (the "phenomenon") from data external to the case (the "context").

Exercise 2.1 Defining the Boundaries of a Case

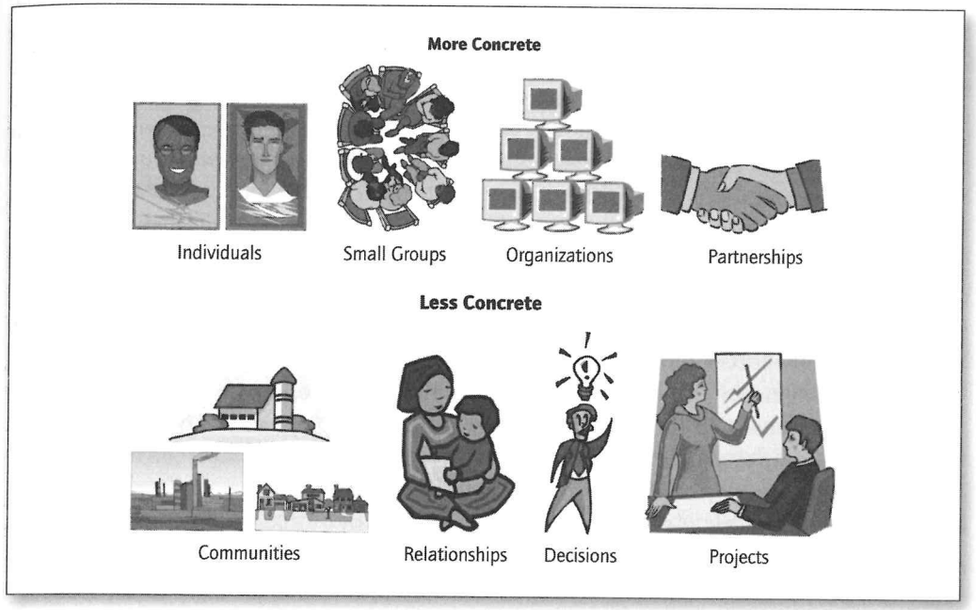
Select a topic for a case study you would like to do. Identify some research questions to be answered or propositions to be examined by your case study. Does the naming of these questions or propositions clarify the boundaries of your case with regard to the time period covered by the case study; the relevant social group, organization, or geographic area; the type of evidence to be collected; and the priorities for data collection and analysis? If not, should you sharpen the original questions?

These latter cautions regarding the need for spatial, temporal, and other concrete boundaries underlie a key but subtle aspect in defining your case. The desired case should be some real-life phenomenon that has some concrete manifestation. The case cannot simply be an abstraction, such as a claim, an argument, or even a hypothesis. These abstractions could rightfully serve as the starting points for research studies using other kinds of methods and not just case study research. To justify doing case study research, you need to go one step further: You need to define a specific, real-life "case" to be a concrete manifestation of the abstraction. (For examples of more concrete and less concrete case study topics, see Figure 2.1.)

Take the concept of "neighboring." Alone, it could be the subject of research studies using methods other than the case study method. The other methods might include a survey of the relationships among neighbors, a history of the evolution of the sense of neighboring and the creation of neighborhood boundaries, or an experiment in which young children do tasks next to each other to determine the distracting effects, if any, of their "neighbors" in a classroom. These examples show how the abstract concept of "neighboring" does not alone produce the grounds for a case study. However, the concept could readily become a case study topic if it were accompanied by your selecting a specific neighborhood ("case") to be studied and posing study questions and propositions about the neighborhood in relation to the concept of "neighboring."

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

Figure 2.1 Illustrative Cases for Case Studies



SOURCE: Clip Art © Jupiter Images.

Linking data to propositions. The fourth component has been increasingly better developed in doing case study research. The component foreshadows the data analysis steps in your case study. Chapter 5 covers these steps and the various analytic techniques and choices in detail. However, during the design stage, you need to be aware of the choices and how they might suit your case study. In this way, your research design can create a more solid foundation for the later analysis.

Exercise 2.2 Defining the Unit of Analysis (and the "Case") for a Case Study

Examine Figure 2.1. Discuss each subject, which illustrates a different unit of analysis. Find a published case study on at least one of these subjects, indicating the actual "case" that was being studied. Understanding that each subject illustrates a different unit of analysis and involves the selection of different cases to be studied, do you think that the more concrete units might be easier to define than the less concrete ones? Why?

All of the analytic techniques in Chapter 5 represent ways of *linking data to propositions:* pattern matching, explanation building, time-series analysis, logic models, and cross-case synthesis. The actual analyses will require that you combine or assemble your case study data as a direct reflection of your initial study propositions. For instance, knowing that some or all of your propositions cover a temporal sequence would mean that you might eventually use some type of time-series analysis. If you noted this strong likelihood during the design phase, you might make sure that your planned data collection included the collection of appropriate time markers as part of the case being studied.

As a caution, if you have had limited experience in conducting empirical studies, you may not easily identify the likely analytic technique(s) or anticipate the needed data to use the techniques to their full advantage. Even more experienced researchers often note how they have either (a) collected too much data that were not later used in any analysis or (b) collected too little data that prevented the proper use of a desired analytic technique. Sometimes, the latter situation even may force researchers to return to their data collection phase (if they can), to supplement the original data. The more you can avoid either of these situations, the better off you will be.

Criteria for interpreting a case study's findings. For many studies, a common illustration of this fifth component arises when statistical analyses are relevant. For instance, by convention, quantitative studies consider a p level of less than .05 to demonstrate that observed differences are "statistically significant" and therefore associated with more robust findings. In other words, the statistical estimates serve as the criteria for interpreting the findings. However, much case study analysis will not rely on the use of statistics, leading to the need to find other ways of thinking about such criteria.

When doing case studies, a major and important alternative strategy is to identify and address rival explanations for your findings. Addressing such rivals becomes a criterion for interpreting your findings: The more rivals that have been addressed and rejected, the stronger will be your findings. Again, Chapter 5 discusses this strategy and how it works. At the design stage of your work, the challenge is to anticipate and enumerate the important rivals, so you will include data about them as part of your data collection. If you only think of rival explanations after data collection has been completed, you will be starting to justify and design a *future* study, but you will not be helping to complete your *current* case study. For this reason, specifying important rival explanations is a part of a case study's research design work.

Summary. A research design should include five components. The first three components—that is, defining your study's questions, propositions, and unit of analysis—will lead your research design into identifying the data that are to be collected. The last two components—that is, defining the logic linking the data to the propositions and the criteria for interpreting the findings—will lead the design into anticipating your case study analysis, suggesting what is to be done after the data have been collected.

THE ROLE OF THEORY OR THEORETICAL PROPOSITIONS IN RESEARCH DESIGNS

Covering the preceding five components of research designs will effectively force you to begin constructing some preliminary theory or theoretical propositions 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 study research and related qualitative methods such as ethnography (Lincoln & Guba, 1985; Van Maanen, 1988) and *grounded theory* (Corbin & Strauss, 2007). Typically, these related methods may deliberately avoid specifying any theoretical propositions at the outset of an inquiry (nor do these methods have to cope with the challenge of defining a "case"). As a result, students who may consider these methods to be interchangeable with case study research wrongly think that, by having selected the case study method, they can proceed quickly into their fieldwork (e.g., by rushing to establish their field contacts as quickly as possible). No presumption 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, some theory development as part of the design phase is highly desired. The needed theory can be plain and simple. For example, a case study on the implementation of a new management information system (MIS) started with the following straightforward theoretical statement:

> 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.

The same MIS case study then added the following theoretical 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 main requirement for implementation to succeed.

You can see that as these two initial statements 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, usefully noted by Sutton and Staw (1995) as "a [hypothetical] story about why acts, events, structure, and thoughts occur" (p. 378). The theoretical propositions can represent key issues from the research literature or practical matters such as differing types of instructional leadership styles or partnering arrangements in a study of organizations. Such propositions will enable the complete research design to provide surprisingly strong guidance in determining the data to collect and the strategies for analyzing the data. For this reason, some theory development prior to the collection of any case study data is desirable. Paul Rosenbaum notes that, for nonexperimental studies more generally, the preferred theoretical statements should elaborate a complex pattern of expected results—the more complex the better (Rosenbaum, 2002, pp. 5–6 and 277–279). The benefit is a stronger design and a heightened ability to interpret your eventual data.

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 (1986) "The Changed World Economy" cited earlier 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 2A in Chapter 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 theoretical statements, and any new empirical study is likely to assume the characteristic of an "exploratory" study. Nevertheless, as noted earlier with the illustrative case in BOX 5, even an exploratory case study should be preceded by statements about what is to be explored, the purpose of the exploration, and the criteria by which the exploration will be judged successful.

Overall, you may want to gain a richer understanding of how theory is used in case studies by reviewing specific case studies that have been successfully completed. For instance, Yin (2012, chap. 3) shows how theory was used in exploratory, descriptive, and explanatory situations by discussing five actual case studies.

Illustrative Topics for 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 (e.g., see H. M. 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 earlier MIS example illustrated 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:

- 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 relations, and interpersonal networks;
- organizational theories—for example, theories of bureaucracies, organizational structure and functions, excellence in organizational performance, and interorganizational partnerships; and
- societal theories—for example, theories of urban development, international conflicts, cultural institutions, technological development, and marketplace functions.

Other examples cut across 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 study research 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. In this situation, Bickman (1987) reminds us that the theory needs to distinguish between the substance of the program (e.g., how to make education more effective) and the process of program implementation (e.g., how to install an effective program). The distinction would avoid situations where policy makers might want to know the desired substantive remedies (e.g., findings about a newly effective curriculum) but where an evaluation unfortunately focused on managerial issues (e.g., the need to hire a good project director). Such a mismatch can be avoided by giving closer attention to the substantive theory.

Use of Theory to Generalize From Case Studies

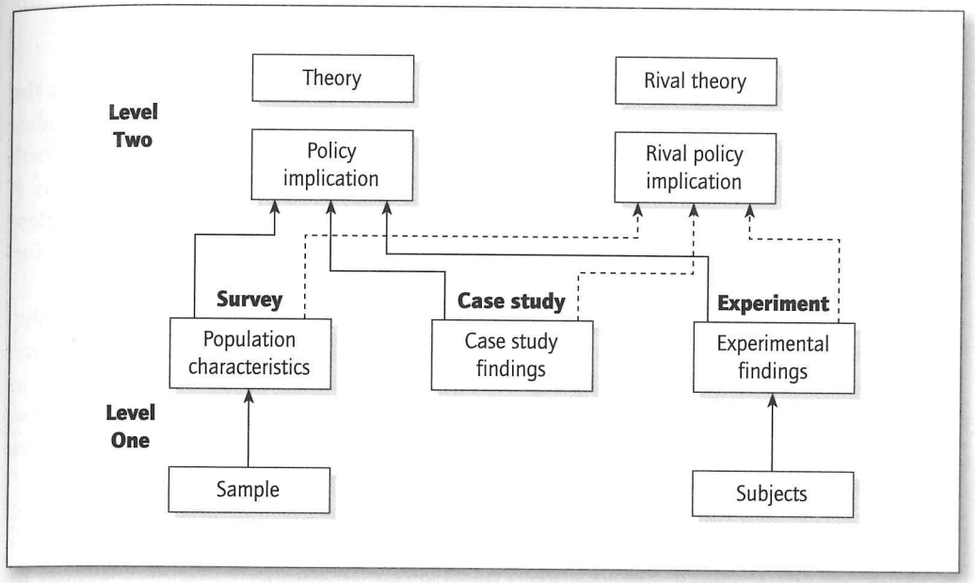
Besides making it easier to design your case study, having some theory or theoretical propositions will later play a critical role in helping you to generalize the lessons learned from your case study. This role of theory has been characterized throughout this book as *analytic generalization* and has been contrasted with another way of generalizing the results from empirical studies, known as *statistical generalization*. Understanding the distinction between these two types of generalization may be your most notable accomplishment in doing case study research.

Let us first take the more commonly recognized way of generalizing—*statistical* generalization—although it is the less relevant one for doing case study research. In statistical generalization, an inference is made about a population (or universe) on the basis of empirical data collected from a sample of that universe. This is shown graphically as a Level One inference in Figure 2.2.[1] This method of generalizing is commonly followed when doing surveys (e.g., Fowler, 1988; Lavrakas, 1987) or analyzing archival data such as in studying housing or employment trends. As another example, political polls need to generalize their findings beyond their sample of respondents and to apply to the larger population, and research investigators readily follow quantitative procedures to determine the confidence with which such extrapolations can be made.

A fatal flaw in doing case studies is to consider statistical generalization to be the way of generalizing the findings from your case study. This is because your case or cases are not "sampling units" and also will be too small in number to serve as an adequately sized sample to represent any larger population.

Rather than thinking about your case as a sample, you should think of it as the opportunity to shed empirical light about some theoretical concepts or principles, not unlike the motive of a laboratory investigator in conceiving of and then conducting a new experiment.[2] In this sense, both a case study and an experiment may have an interest in going beyond the specific case or experiment. Both kinds of studies are likely to strive for generalizable findings or lessons learned—that is, analytic generalizations—that go beyond the setting for the specific case or specific experiment that had been studied (see Tutorial 2-1). For example, the lessons learned could assume the form of a *working hypothesis*

Figure 2.2 Making Inferences: Two Levels



(Cronbach, 1975), either to be applied in reinterpreting the results of existing studies of other concrete situations (that is, other cases or experiments) or to define new research focusing on yet additional concrete situations (that is, new cases or experiments). Note that the aim of an analytic generalization is still to generalize to these other concrete situations and not just to contribute to abstract theory building. Also note that the generalizations, principles, or lessons learned from a case study may potentially apply to a variety of situations, far beyond any strict definition of the hypothetical population of "like-cases" represented by the original case.

The theory or theoretical propositions that went into the initial design of your case study, as empirically enhanced by your case study's findings, will have formed the groundwork for an analytic generalization. Alternatively, a new generalization may emerge from the case study's findings alone. In other words, the *analytic generalization* may be based on either (a) corroborating, modifying, rejecting, or otherwise advancing theoretical concepts that you referenced in designing your case study or (b) new concepts that arose upon the completion of your case study. The important point is that, regardless of whether the generalization was derived from the conditions you specified at the outset or uncovered at the conclusion of your case study, the generalization will be at a conceptual level higher than that of the specific case (or experiment)—shown graphically as a Level Two inference in Figure 2.2.

Several prominent case studies illustrate how analytic generalizations can use a case study's findings to implicate new situations. First, consider how the three initial case studies of this book (cited in BOXES 1, 2A, and 2B of Chapter 1) treated the generalizing function:

- **BOX 1:** Allison's case (1971) is about the Cuban missile crisis, and he relates the three theoretical models from his case study to many other situations, first to other international confrontations, such as between the United States and North Vietnam in the 1960s (p. 258). The later edition of the case study (Allison & Zelikow, 1999) then discusses the models' relevance to the "rethinking of nuclear threats to Americans today" (p. 397) as well as to the broader challenge of inferring the motives underlying actions taken by a foreign power.
- **BOX 2A:** Whyte's study (1943/1993) is well known for uncovering the relationship between individual performance and group structure, highlighted by a bowling tournament where he directly experienced the impact on his own performance ("as if something larger than myself was controlling the ball"—p. 319) and observed how the gang members' bowling scores, with one notable exception, emulated their standing in the gang. Whyte generalizes his findings by later commenting that "I believed then (and still believe now) that this sort of relationship may be observed in other group activities everywhere" (p. 319).
- **BOX 2B:** Neustadt and Fineberg (1983) show yet another variation, claiming no generalization but concluding with an extensive discussion of the usefulness of their case study in teaching public policy courses (pp. 231–250).

Second, BOX 7 contains four additional illustrations. All show how findings from a single-case study nevertheless can be generalized to a broad variety of other situations. The fourth of these case studies has one other notable feature: It demonstrates how an entire case study can be published as a journal article (the first three examples appeared in the form of rather lengthy books).

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. Also to come, this chapter's later discussion under the topic of *external validity* adds a further insight about making analytic generalizations. The main point at this juncture is that you should try to aim toward analytic generalizations 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-or multiple-case study were equivalent to respondents in a survey. In other words, again as graphically depicted in Figure 2.2, you should aim for Level Two inferences when doing case studies.

In a like manner, even referring to your case or cases as a "purposive sample" may raise similar conceptual and terminological problems. You may have intended to convey that the "purposive" portion of the term reflects your selection of a case that will illuminate the theoretical propositions of your case study. However, your use of the

BOX 7
Generalizing from Single-Case Studies: Four More Examples

7A. A Sociology of "Mistake"

The tragic loss of the space shuttle *Challenger* in 1986, vividly shown in repeated TV replays of the spaceship's final seconds, certainly qualifies as a unique case. The causes of this loss became the subject of a Presidential Commission and of a case study by Diane Vaughan (1996). Vaughan's detailed study shows how the social structure of an organization (the NASA space agency) had, over time, transformed deviance into acceptable and routine behavior.

Vaughan's ultimate explanation differs markedly from that of the Presidential Commission, which pointed to individual errors by middle managers as the main reasons for failure. In Vaughan's words, her study "explicates the sociology of mistake" (p. xiv)–that "mistakes are systemic and socially organized, built into the nature of professions, organizations, cultures, and structures" (p. 415). She shows how deviance is transformed into acceptable behavior through the institutionalization of production pressures (originating in the organizational environment), leading to "nuanced, unacknowledged, pervasive effects on decisionmaking" (p. xiv). Her final discussion applies this generalization to a diverse array of other situations. As examples, she cites studies showing the research distortions created by the worldview of scientists, the uncoupling of intimate relationships, and the inevitability of accidents in certain technological systems.

7B. The Origins of Social Class

The second example is about the uncovering and labeling of a social class structure based on a case study of a small American city, Yankee City (Warner & Lunt, 1941). This classic case study in sociology made a critical contribution to social stratification theory and an understanding of the social differences among "upper," "upper-middle," "middle-middle," "upper-lower," and "lower" classes. Over the years, the insights from these differences have applied to a broad range of situations (by no means limited to other small cities).

7C. Contribution to Urban Planning

The third example is Jane Jacobs and her famous book, *The Death and Life of Great American Cities* (1961). The book is based mostly on experiences from a single case, New York City. The book's chapters then show how these New York experiences can be used to develop broader theoretical principles 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.

(Continued)

(Continued)

Jacobs's book created heated controversy in the planning profession. New empirical inquiries were made about one or another of her rich and provocative ideas. These inquiries helped to test the broader applicability of her principles to other concrete settings, and in this way Jacobs's work still stands as a significant contribution in the field of urban planning.

7D. Government Management of "Spoiled" National Identity

The fourth example creatively extended Erving Goffman's well-known sociological theory, regarding the management of stigma by individual people, to an institutional level (Rivera, 2008). A field-based case study of Croatia showed how the stigma created by the wars of Yugoslav secession had demolished the country's image as a desirable tourist destination, but then how the country successfully used an impression management strategy to revive the tourism. Croatia thus presented "an exciting case of reputation management in action" (p. 618). The author suggests that her adapted theoretical model can be used as "a launching point for understanding the public representation dilemmas faced by other states and organizational actors that have undergone reputation-damaging events" (p. 615). In so doing, the case study has provided another illustration of analytic generalization.

"sample" portion of the term still risks misleading others into thinking that the case comes from some larger universe or population of like-cases, undesirably reigniting the specter of statistical generalization. The most desirable posture may be to avoid referring to any kind of sample (purposive or otherwise). (The preferred criteria and terminology for selecting cases, as part of either a single- or a multiple-case study, are discussed later in this chapter under the topic of "case study designs.") In this sense, case study research directly parallels experimental research: Few if any people would consider that a new experiment should be designed as a sample (of any kind) from a larger population of like-experiments—and few would consider that the main way of generalizing the findings from a single experiment would be in reference to a population of like-experiments.

Summary

This section has suggested that a complete research design, while including the five components previously described, will benefit from the development of theoretical propositions. A good case study researcher should pursue such propositions and take advantage of this benefit, whether the case study is to be exploratory, descriptive, or explanatory. The use of theory and theoretical propositions in doing case studies is an immense aid in defining the appropriate research design and data to be collected. The

same theoretical orientation also will become the main vehicle for generalizing the findings from 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. Government Accountability Office, 1990).

Four tests, however, have been commonly used to establish the quality of any empirical social research. Because case study research is part of this larger body, the four tests also are relevant to case study research. An important innovation of this book is the identification of several tactics for dealing with these four tests when doing case study research. 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 described in detail in the referenced chapter of this book.)

Because the four tests are common to all social science methods, the tests have been summarized in numerous textbooks (e.g., see Kidder & Judd, 1986, pp. 26–29). The tests also have served as a framework for assessing a large group of case studies in the field of strategic management (Gibbert, Ruigrok, & Wicki, 2008). The four tests are:

Figure 2.3 Case Study Tactics for Four Design Tests

TESTS	Case Study Tactic	Phase of Research in which Tactic Occurs
Construct validity	• use multiple sources of evidence • establish chain of evidence • have key informants review draft case study report	data collection (see Chap. 4) data collection (see Chap. 4) composition (see Chap. 6)
Internal validity	• do pattern matching • do explanation building • address rival explanations • use logic models	data analysis (see Chap. 5) data analysis (see Chap. 5) data analysis (see Chap. 5) data analysis (see Chap. 5)
External validity	• use theory in single-case studies • use replication logic in multiple-case studies	research design (see Chap. 2) research design (see Chap. 2)
Reliability	• use case study protocol • develop case study database	data collection (see Chap. 3) data collection (see Chap. 4)

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

Each item on this list deserves explicit attention. For case study research, an important revelation is that the several tactics to be used in dealing with these tests should be applied throughout the subsequent conduct of a case study, not just at its beginning. Thus, the "design work" for doing case studies may actually continue beyond the initial design plans.

Construct Validity

This first test is especially challenging in case study research. People who have been critical of case studies often point to the fact that a case study researcher fails to develop a sufficiently operational set of measures and that "subjective" judgments—ones tending to confirm a researcher's preconceived notions (Flyvberg, 2006; Ruddin, 2006)—are used to collect the data.[3] Take an example such as studying "neighborhood change"—a common case study topic (e.g., Bradshaw, 1999; Keating & Krumholz, 1999): Over the years, concerns have arisen over how certain urban neighborhoods have changed their character. Any number of case studies has 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 claimed changes in a case study genuinely reflect the events in a neighborhood or whether they happen to be based on a researcher's impressions only.

Neighborhood change can 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 revitalizing neighborhoods. The choice of whether to aggregate blocks, census tracts, or larger areas also can produce different results (Hipp, 2007).

To meet the test of construct validity, an investigator must be sure to cover two steps:

1. define neighborhood change in terms of specific concepts (and relate them to the original objectives of the study) and
2. identify operational measures that match the concepts (preferably citing published studies that make the same matches).

For example, suppose you satisfy the first step by stating that you plan to study neighborhood change by focusing on trends in neighborhood crime. The second step now demands that you select a specific measure, such as police-reported crime (which happens to be the standard measure used in the FBI Uniform Crime Reports) as your measure of crime. The literature will indicate certain known shortcomings in this measure, mainly that unknown proportions of crimes are not reported to the police. You will then need to discuss how the shortcomings nevertheless will not bias your study of neighborhood crime and hence neighborhood change.

As previously shown in Figure 2.3, three tactics are available to increase construct validity when doing case studies. 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 (also 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 internal 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 mainly a concern for explanatory case studies, when an investigator is trying to explain how and why 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 this kind of causal situation.

Second, the concern over internal validity, for case study research, extends to the broader problem of making inferences. Basically, a case study involves an inference every time an event cannot be directly observed. 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 when doing case study research. Figure 2.3 (previously shown) suggests four analytic tactics.

All are described further in Chapter 5 because they take place during the analytic phase of doing case studies: *pattern matching, explanation building, addressing rival explanations,* and *using logic models*.

External Validity

The third test deals with the problem of knowing whether a study's findings are generalizable beyond the immediate study, regardless of the research method used (e.g., experiments, surveys, or case studies). For case studies, the issue relates directly to the earlier discussion of *analytic generalization* and the reference to Level Two in Figure 2.2. To repeat the earlier discussion further, referring to *statistical generalization* and any analogy to samples and populations would be misguided.

Another insight on this issue derives from observing the form of the original research question(s) posed in doing your case study. The form of the question(s) can help or hinder the preference for seeking generalizations—that is, striving for external validity.

Recall that the decision to favor case study research should have started with the posing of some "how" and "why" question(s). For instance, many descriptive case studies deal with the "how" of a situation, whereas many explanatory case studies deal with the "why" of situations. However, if a case study has no pressing "how" or "why" questions—such as a study merely wanting to document the social trends in a neighborhood, city, or country or the employment trends in an organization (and essentially posing a "what" question)—arriving at an analytic generalization may be more difficult. To avoid this situation, augmenting the study design with "how" and "why" questions (and collecting the additional data) can be extremely helpful. (Alternatively, if in the illustrative examples a study's research interest is entirely limited to documenting social trends, using some other method might serve the study's objectives better than using the case study method.)

In this manner, the form of the initial research question(s) can directly influence the strategies used in striving for external validity. These research question(s) should have been settled during the research design phase of your case study, if not earlier. For this reason, Figure 2.3 as previously shown points to the research design phase, with the identification of appropriate theory or theoretical propositions, as being the most appropriate time for establishing the groundwork for starting to address the external validity of your case study.

Reliability

Most people are probably already familiar with this final test. The objective is to be sure that, if a later researcher follows the same procedures as described by an earlier researcher and conducts the same case study 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 (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 method.[4] As previously shown, Figure 2.3 suggests two specific tactics for overcoming these shortcomings—the use of a *case study protocol* to deal with the documentation problem in detail (discussed in Chapter 3) and the development of a *case study database* (discussed in Chapter 4).

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 looking over your shoulder. Accountants and bookkeepers always are aware that any calculations must be capable of being audited. In this sense, an auditor also is 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 in principle repeat the procedures and hopefully 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 same phase in doing 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.

Exercise 2.3 Defining the Criteria for Judging the Quality of Research Designs

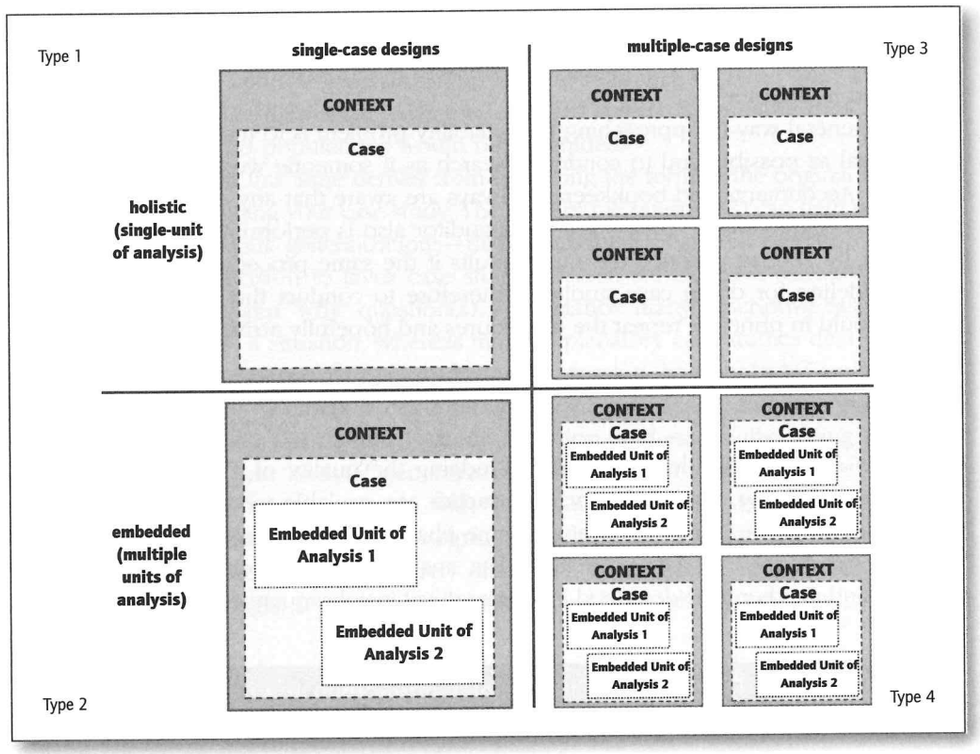
Define the four criteria for judging the quality of research designs: (a) construct validity, (b) internal validity, (c) external validity, and (d) reliability. Give an example of each type of criterion in a case study you might want to do.

CASE STUDY RESEARCH DESIGNS

Traditional case study research has not usually included the idea of having formal designs, as might be found when doing survey or experimental research. You still may successfully

conduct a new case study without any formal design. However, attending to the potential case study research designs can make your case studies stronger and, possibly, easier to do. You might therefore find the remainder of this section to be useful. It covers four types of designs, based on the 2 × 2 matrix in Figure 2.4.

Figure 2.4 Basic Types of Designs for Case Studies



SOURCE: COSMOS Corporation.

The matrix first shows that every type of design will include the desire to analyze contextual conditions in relation to the "case," with the dotted lines between the two signaling that the boundaries between the case and the context are not likely to be sharp. The matrix then shows that single- and multiple-case studies reflect different design situations and that, within these two variants, there also can be unitary or multiple units of analysis. The resulting four types of designs for case studies are (Type 1) single-case (holistic) designs, (Type 2) single-case (embedded) designs, (Type 3) multiple-case (holistic) designs, and (Type 4) multiple-case (embedded) designs. The rationale for these four types of designs is as follows.

What Are the Potential Single-Case Designs (Types 1 and 2)?

Five rationales for single-case designs. A primary distinction in designing case studies is between *single-* and *multiple-*case study designs. This means the need for a decision, prior to any data collection, on whether you are going to use a single case or multiple cases in your case study.

The single-case study is an appropriate design under several circumstances, and five single-case rationales—that is, having a *critical, unusual, common, revelatory,* or *longitudinal* case—are given below. Recall that a single-case study is analogous to a single experiment, and many of the same conditions that justify a single experiment also can justify a single-case study.

Recall, too, that the selection of your case should be related to your theory or theoretical propositions of interest. These form the substantive context for each of the five rationales. Thus, the first rationale for a single case—selecting a *critical case*—would be critical to your theory or theoretical propositions (again, note the analogy to the critical *experiment*). The theory should have specified a clear set of circumstances within which its propositions are believed to be true. The single case then can be used to determine whether the 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 and the Cuban missile crisis (described in Chapter 1, BOX 2), the single case can represent a significant contribution to knowledge and theory building by confirming, challenging, or extending the theory. Such a study even can help to refocus future investigations in an entire field. (See BOX 8 for another example, in the field of organizational innovation.)

BOX 8
The Critical Case as a Single-Case Study

One rationale for selecting a single-case rather than a multiple-case design is that the single case can represent the critical test of a significant theory. Gross, Bernstein, and Giacquinta (1971) used such a design by focusing on a single school in their book, *Implementing Organizational Innovations* (also see BOX 20B, p. 114).

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. (1971) 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 failure.

In this manner, the book, though limited to a single case, represented a watershed in organizational innovation theory. Prior to the study, analysts had focused on the identification of barriers to innovation; since the study, the literature has been much more dominated by studies of the implementation process, not only in schools but also in many other types of organizations.

A second rationale for a single case is where the case represents an *extreme* case or an *unusual* case, deviating from theoretical norms or even everyday occurrences. For instance, such cases can occur in clinical psychology, where a specific injury or disorder may offer a distinct opportunity worth documenting and analyzing. In clinical research, a common research strategy calls for studying these unusual cases because the findings may reveal insights about normal processes. In this manner, the value of a case study can be connected to a large number of people, well beyond those suffering from the original clinical syndrome.

Conversely, a third rationale for a single case is the *common* case. Here, the objective is to capture the circumstances and conditions of an everyday situation—again because of the lessons it might provide about the social processes related to some theoretical interest. In this manner, a street scene and its sidewalk vendors can become the setting for learning about the potential social benefits created by informal entrepreneurial activity (e.g., Duneier, 1999), a study of a small business can yield insights into innovations and innovative processes (e.g., see Yin, 2012, chap. 9), and the social and institutional structure within a single, low-income urban neighborhood can provide insights into the relationship between poverty and social capital (e.g., Small, 2004).

A fourth rationale for a single-case study is the *revelatory* case. This situation exists when a researcher has an opportunity to observe and analyze a phenomenon previously inaccessible to social science inquiry, such as Whyte's (1943/1993) *Street Corner Society,* previously described in Chapter 1, BOX 2A. Another example is Phillippe Bourgois's (2003) study of crack and the drug-dealing marketplace in Spanish Harlem—a neighborhood in New York City. The author gained the trust and long-term friendship of two dozen street dealers and their families, revealing a lifestyle that few had been able to study up to that time. For another example, see Elliot Liebow's (1967) famous case study of unemployed men, *Tally's Corner* (BOX 9). When researchers have similar types of opportunities and can uncover some prevalent phenomenon previously inaccessible to social scientists, such conditions justify the use of a single-case study on the grounds of its revelatory nature.

BOX 9
The Revelatory Case as a Single-Case Study

Another rationale for selecting a single case is that the researcher has access to a situation previously inaccessible to empirical study. The case study is therefore worth conducting because the descriptive information alone will be revelatory.

Such was the situation in Elliot Liebow's (1967) sociological classic, *Tally's Corner.* The book is about a single group of African American men living in a poor, inner-city neighborhood. By befriending these men, the author was able to learn about their lifestyles, their

(Continued)

(Continued)

coping behavior, and in particular their sensitivity to unemployment and failure. The book provided insights into socioeconomic conditions that have prevailed in many U.S. cities for a long time, but that only had been obscurely understood. The single case showed how investigations of such topics could be done, thus stimulating much further research and eventually the development of needed public policy actions.

A fifth rationale for a single-case study is the *longitudinal* case: studying the same single case at two or more different points in time. The theory of interest would likely specify how certain conditions and their underlying processes change over time. The desired time intervals would presumably reflect the anticipated stages at which the changes should reveal themselves. They may be pre-specified time intervals, such as prior to and then after some critical event, following a "before" and "after" logic. Alternatively, they might not deal with specific time intervals but cover trends over an elongated period of time, following a developmental course of interest. Under exceptional circumstances, the same case might be the subject of two consecutive case studies, such as occurred with *Middletown* (Lynd & Lynd, 1929) and *Middletown in Transition* (Lynd & Lynd, 1937). Whatever the time intervals or periods of interest, the processes being studied should nevertheless reflect the theoretical propositions posed by the case study. The desired time intervals would presumably reflect the anticipated stages at which the changes should reveal themselves, and the processes being observed should again reflect the theoretical propositions posed by the case study.

These five serve as major rationales for selecting a single-case study. There are other situations in which the single-case study may be used as a pilot case that might be the beginning of a multiple-case study. However, in this latter situation, the single-case portion of the study would not be regarded as a complete study on its own.

Whatever the rationale for doing single-case studies (and there may be more than the five 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 yourself to any single-case study until these major concerns have been covered.

Holistic versus embedded case studies. The same single-case study may involve units of analysis at more than one level. This occurs when, within a single case, attention is also given to a subunit or subunits (see BOX 10). For instance, even though a case

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 "Kinds of Data" table, below). 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.

Kinds of Data					
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*
International Typographical Union 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	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, and Coleman (1956, p. 422). Reprinted by permission.

study might be about a single organization, such as a hospital, the analysis might include outcomes about the clinical services and staff employed by the hospital (and possibly even some quantitative analyses based on the employee records of the staff). In an evaluation study, the single case might be a public program that involves large numbers of funded projects—which would then be the embedded units (see Appendix B for more details). In either situation, these embedded units can be selected through sampling or cluster techniques (McClintock, 1985). No matter how the units are selected, the resulting design 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 an organization or of a program, a *holistic design* would have been used (see Figure 2.4, Type 1).

These two variants of single-case studies both have their strengths and weaknesses. The holistic design is advantageous when no logical subunits can be identified or when the relevant theory underlying the case study is itself of a holistic nature. Potential problems arise, however, when a global approach allows a researcher to avoid examining any specific phenomenon in operational detail. Thus, a typical problem with the holistic design is that the entire case study may be conducted at an unduly abstract level, lacking sufficiently 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 research questions. Although some people have claimed such flexibility to be a strength of case study research, in fact the largest criticism of case studies is based on this type of shift—in which the implemented research design is no longer appropriate for the research questions being asked (see COSMOS Corporation, 1983). Because of this problem, you need to avoid such unsuspected slippage; 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. Thus, an embedded design can serve as an important device for focusing a case study inquiry.

An embedded design, however, also has its 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. For instance, an evaluation of a program consisting of multiple projects may include project characteristics as a subunit of analysis. The project-level data may even be highly quantitative if there are many projects. However, the original evaluation becomes a project study (i.e., a multiple-case study of different projects) if no investigating is done at the level of the original case—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 employee and not an organizational study. In both examples, what has happened is that

the original phenomenon of interest (a program or organizational climate) has become the context and not the target of study.

Summary. Single-case studies are a common design for doing case study research, and two variants 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) a critical test of existing theory, (b) an extreme or unusual circumstance, or (c) a common case, or where the case serves a (d) revelatory or (e) longitudinal 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 caution must be exercised—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 study 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. If the shift is justifiable, you need to address it explicitly and indicate its relationship to the original inquiry.

What Are the Potential Multiple-Case Designs (Types 3 and 4)?

The same study may contain more than a single case. When this occurs, the study has used 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 the use of new curricula, rearranged school schedules, or a new educational technology), in which individual schools adopt some innovation. Each school might be the subject of an individual case study, but the study as a whole covers several schools and in this way uses a multiple-case design.

Multiple- versus single-case designs. In some fields, multiple-case studies have been considered a different "methodology" from 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; Lijphart, 1975). This book, however, considers single- and multiple-case designs to be variants 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 case study research.

Multiple-case designs have distinct advantages and disadvantages in comparison to 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 cannot usually be satisfied by multiple cases. By definition, the unusual or extreme case, the critical case, and the revelatory case all are likely to involve only single cases. 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.

Selecting the multiple cases also raises a new set of questions. Here, *a major insight is to consider multiple cases as one would consider multiple experiments*—that is, to follow a "replication" design. This is far different from a misleading analogy that incorrectly considers 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" design. The methodological differences between these two views are revealed by the different rationales underlying the replication as opposed to sampling designs.

Replication, not sampling logic, for multiple-case studies. The replication logic is analogous to that used in multiple experiments (see Hersen & Barlow, 1976). For example, upon uncovering a significant finding from a single experiment, an ensuing and pressing priority would be to replicate this finding by conducting a second, third, and even more experiments. Some of the replications might attempt to duplicate the exact conditions of the original experiment. Other replications might alter one or two experimental conditions considered unimportant to the original finding, to see whether the finding could still be duplicated. Only with such replications would the original finding be considered robust.

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) predicts contrasting results but for anticipatable reasons (a *theoretical replication*). The ability to conduct 6 or 10 case studies, arranged effectively within a multiple-case design, is analogous to the ability to conduct 6 to 10 experiments on related topics; a few cases (2 or 3) would be literal replications, whereas a few other cases (4 to 6) might be designed to pursue two different patterns of theoretical replications. If all the cases turn out as predicted, these 6 to 10 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 researchers deal with conflicting experimental findings.

The logic underlying these replication procedures also should reflect some theoretical interest, not just a prediction that two cases should simply be similar or different. For example, one might consider the initial proposition that an increase in using a new

computer system in small business environments will occur when the system is used for both administrative (e.g., accounting and personnel) and business (e.g., sales and production) applications, but not either alone. To pursue this proposition in a multiple-case study design, 3 or 4 small businesses (or cases) might be selected in which both types of applications are present, to determine whether, in fact, use of the system did increase over a period of time (the investigation would be predicting a literal replication in these 3 or 4 cases). Three or 4 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, 3 or 4 other cases would be selected in which only business applications 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 9 to 12 cases, in the aggregate, would provide substantial support for the initial proposition.

Another example of a multiple-case replication design comes from the field of urban studies (see BOX 11). You also can find examples of three entire case studies, all following a replication design but covering university administration, the transformation of business firms, and HIV/AIDS prevention, in the companion text (Yin, 2012, chaps. 11, 12, and 15).

BOX 11
A Multiple-Case, Replication Design

A common problem in the 1960s and 1970s was how to get good advice to city governments. Peter Szanton's (1981) book, *Not Well Advised,* 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 city governments. 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 but also then putting advice into practice.

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.

This replication logic, whether applied to experiments or to case studies, must be distinguished from the sampling logic commonly used in surveys. The sampling logic requires an operational enumeration of the entire universe or pool of potential respondents and then a statistical procedure for selecting a specific subset of respondents to be surveyed. 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 presumed accurate. The entire procedure is commonly used when a researcher wishes to determine the prevalence or frequency of a particular phenomenon.

Any application of this sampling logic to case study research would be misplaced. First, case studies are not the best method for assessing the prevalence 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 sample of cases—too large to allow more than a superficial examination of any given case.

Third, if a sampling logic had to be applied to all types of research, many important topics could not be empirically investigated, such as the following problem: Your investigation deals with the role of the presidency of the United States, and you are interested in doing a multiple-case study of (a few) presidents to test your theory about presidential leadership. However, the complexity of your topic means that your choice of a small number of cases could not adequately represent all the 44 presidents since the beginning of the Republic. Critics using a sampling logic might therefore deny the acceptability of your study. In contrast, if you use a replication logic, the study is eminently feasible.

The replication approach to multiple-case studies is illustrated in Figure 2.5. 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.

An important part of Figure 2.5 is the dashed-line feedback loop. The loop represents the situation where important discovery occurs during the conduct of one of the individual case studies (e.g., one of the cases did not in fact suit the original design). Such a discovery may require you to reconsider one or more of the study's original theoretical propositions. At this point, "redesign" should take place before proceeding

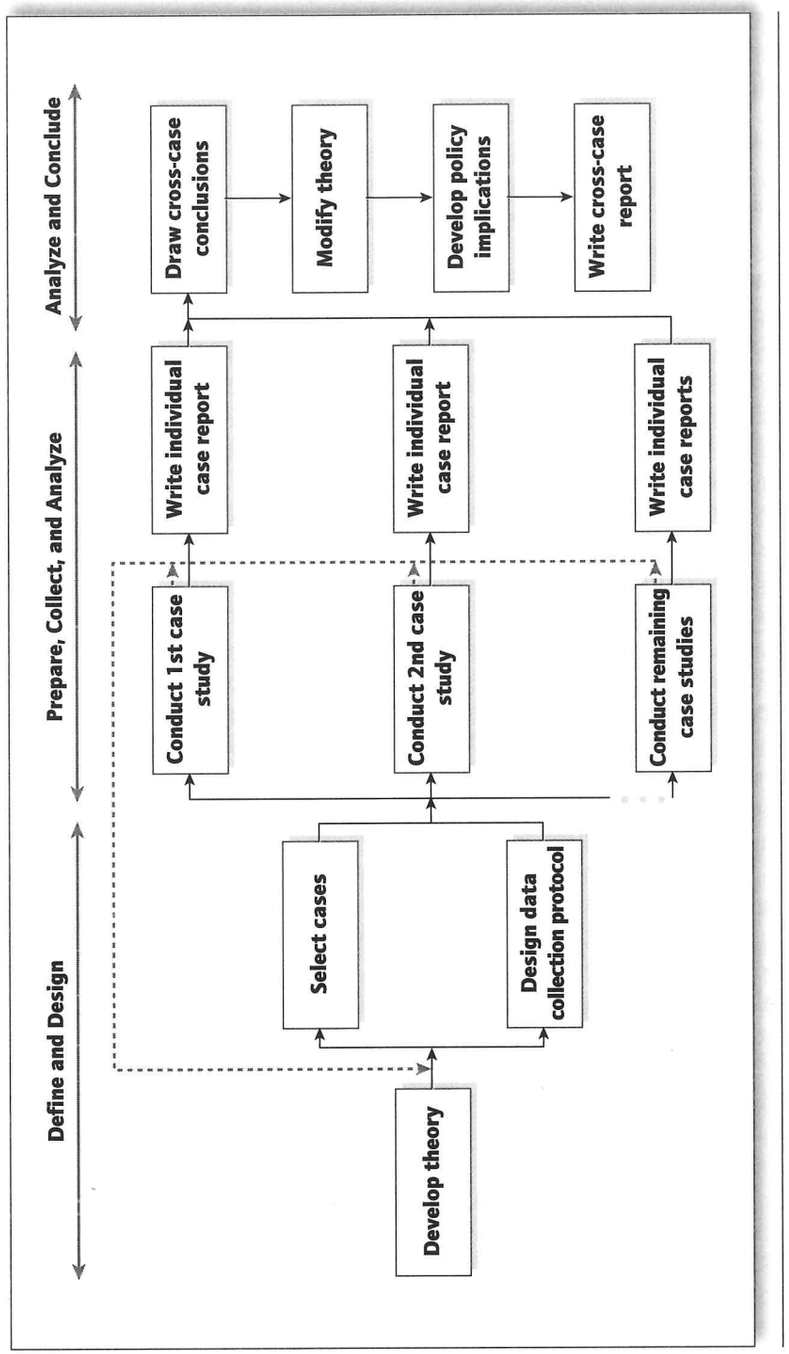


Figure 2.5 Multiple-Case Study Procedure

SOURCE: COSMOS Corporation.

further. Such redesign might involve the selection of alternative cases or changes in the case study protocol (see Chapter 3). Without such redesign, you risk being accused of distorting or ignoring the discovery, just to accommodate the original design. This condition leads quickly to a further accusation—that you have been selective in reporting your data, to suit your preconceived ideas (that is, the original theoretical propositions).

Overall, Figure 2.5 depicts a different logic from that of a sampling design. The logic as well as its contrast with a sampling design may be difficult to follow and is worth extensive discussion with colleagues before proceeding with any multiple-case study.

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 the use of a power analysis to determine the desired sample size (e.g., Lipsey, 1990) also are irrelevant. Instead, you should think of the number of case replications—both literal and theoretical—that you need or would like to have in your study.

Your judgment will be a discretionary, not formulaic, one. Such discretionary judgments occur in non–case study research, such as in setting the criterion for defining a "significant effect" in experimental science. Thus, designating a "$p < .05$" or "$p < .01$" likelihood of detection, to set the confidence level for accepting or rejecting the null hypothesis, is not based on any formula but is a matter of discretionary, judgmental choice. Note that when patient safety and well-being are at stake, as in a clinical trial, investigators will usually not even settle for a "$p < .01$" significance level but may choose to attain a "$p < .0001$" or even more stringent level. Analogously, designating the number of replications depends upon the certainty you want to have about your multiple-case results. For example, you may want to settle for two or three literal replications when your theory is straightforward and the issue at hand does not demand an excessive degree of certainty. However, if your theory is subtle or if you want a higher degree of certainty, you may press for five, six, or more replications.

In deciding upon the number of replications, an important consideration is related to your sense of the strength and importance of rival explanations. The stronger the rivals, the more additional cases you might want, each case showing a different result when some rival explanation had been taken into account. For example, your original hypothesis might be that summer reading programs improve students' reading scores, and you already might have shown this result through several cases that served as literal replications. A rival explanation might be that parents also work more closely with their children during the summer and that this circumstance can account for the improved reading scores. You would then find another case, with parent participation but no summer reading program, and in this theoretical replication you would predict that the scores would not improve. Having two such theoretical replications would provide even greater support for your findings.

Rationale for multiple-case designs. In short, the rationale for multiple-case designs derives directly from your understanding of literal and theoretical replications. The

simplest multiple-case design would be the selection of two or more cases that are believed to be literal replications, such as a set of cases with exemplary outcomes in relation to some evaluation question, such as "how and why a particular intervention has been implemented smoothly." Selecting such cases requires prior knowledge of the outcomes, with the multiple-case inquiry focusing on how and why the exemplary outcomes might have occurred and hoping for literal (or direct) replications of these conditions from case to case.[5]

More complicated multiple-case designs would likely result from the number and types of theoretical replications you might want to cover. For example, investigators have used a "two-tail" design in which cases from both extremes (of some important theoretical condition, such as extremely good and extremely bad outcomes) have been deliberately chosen. Multiple-case rationales also can derive from the prior hypothesizing of different types of conditions and the desire to have subgroups of cases covering each type. These and other similar designs are more complicated because the study should still have at least two individual cases within each of the subgroups, so that the theoretical replications across subgroups are complemented by literal replications within each subgroup.

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-case studies: 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 variants depends upon the type of phenomenon being studied and your research questions. In an embedded design, a study even may call for the conduct of a survey at each case study site. For instance, suppose a study is concerned with the impact of the same type of curriculum adopted by different nursing schools. Each nursing school may be the topic of a case study, with the theoretical framework dictating that nine such schools be included as case studies, three to replicate a direct result (literal replication) and six others to deal with contrasting conditions (theoretical replications).

For all nine schools, an embedded design is used because surveys of the students (or, alternatively, examination of students' archival records) are needed to address research questions about the performance of the schools. However, the results of each survey will *not* be pooled across schools. Rather, the survey data will be part of the findings for each individual nursing school, or case. These data may be highly quantitative and even involve statistical tests, focusing on the attitudes and behavior of individual students, and the data will be used along with information about the school to interpret the success and operations with the curriculum at that particular school. If, in contrast, the survey data are pooled across schools, a replication design is no longer being used. In fact, the study has now become an embedded, single-case study, in which all nine

schools and their students have now become part of some larger, main unit of analysis that might not have been specified at the outset. Such a turn of events would create a pressing need to discard the original multiple-case design. The newly designed single-case study would require a complete redefinition of the main unit of analysis and entail extensive revisions to the original theories and propositions of interest.

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 a researcher 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 individual cases within a multiple-case study design may be either holistic or embedded. When an embedded design is used, each individual case study may in fact include the collection and analysis of quantitative data, including the use of surveys within each case study.

Exercise 2.4 Defining a Case Study Research Design

Select one of the case studies described in the BOXES of this book, reviewing the entire case study (not just the material in the BOX). Describe the research design of this case study. How did it justify the relevant evidence to be sought, given the main research questions to be answered? What methods were used to identify the findings, based on the evidence? Is the design a single- or multiple-case design? Is it holistic or does it have embedded units of analysis?

MODEST ADVICE IN SELECTING CASE STUDY DESIGNS

Now that you know how to define case study designs and are prepared to carry out design work, you might want to consider three pieces of advice.

Single- or Multiple-Case Designs?

The first word of advice is that, although all designs can lead to successful case studies, when you have the choice (and resources), multiple-case designs may be preferred over single-case designs. If you can do even a "two-case" case study, your chances of doing a

good case study will be better than using a single-case design. Single-case designs are vulnerable if only because you will have put "all your eggs in one basket." More important, the analytic benefits from having two (or more) cases may be substantial.

To begin with, even with two cases, you have the possibility of direct replication. Analytic conclusions independently arising from two cases, as with two experiments, will be more powerful than those coming from a single case (or single experiment) alone. Alternatively, you may have deliberately selected your two cases because they offered contrasting situations, and you were not seeking a direct replication. In this design, if the subsequent findings support the hypothesized contrast, the results represent a strong start toward theoretical replication—again strengthening your findings compared to those from a single-case study alone (e.g., Eilbert & Lafronza, 2005; Hanna, 2005; also see BOX 12).

BOX 12
Two, "Two-Case" Case Studies

12A. Contrasting Cases for Community Building

Chaskin (2001) used two case studies to illustrate contrasting strategies for capacity building at the neighborhood level. The author's overall conceptual framework, which was the main topic of inquiry, claimed that there could be two approaches to building community capacity–using a collaborative organization to (a) reinforce existing networks of community organizations or (b) initiate a new organization in the neighborhood. After thoroughly airing the framework on theoretical grounds, the author presents the two case studies, showing the viability of each approach.

12B. Contrasting Strategies for Educational Accountability

In a directly complementary manner, Elmore, Abelmann, and Fuhrman (1997) chose two case studies to illustrate contrasting strategies for designing and implementing educational accountability (i.e., holding schools accountable for the academic performance of their students). One case represented a lower cost, basic version of an accountability system. The other represented a higher cost, more complex version.

In general, criticisms about single-case studies usually reflect fears about the uniqueness or artifactual conditions surrounding the case (e.g., special access to a key informant). As a result, the criticisms may turn into skepticism about your ability to do empirical work beyond having done a single-case study. Having two cases can begin to blunt such criticism and skepticism. Having more than two cases will produce an even stronger effect. In the face of these benefits, having at least two cases should be your goal. If you do use a single-case design, you should be prepared to make an extremely strong argument in justifying your choice for the case.

Exercise 2.5 Establishing the Rationale for a Multiple-Case Study

Develop some preliminary ideas about a "case" for your case study. Alternatively, focus on one of the single-case studies presented in the BOXES in this book. In either situation, now think of a companion "case" that might augment the single case. In what ways might the companion case's findings supplement those of the first case? Could the data from the second case fill a gap left by the first case or respond better to some obvious shortcoming or criticism of the first case? Would the two cases together comprise a stronger case study? Could yet a third case make the findings even more compelling?

Closed or Adaptive Designs?

Another word of advice is that, despite this chapter's details about design choices, you should not think that a case study's design cannot be modified by new information or discovery during data collection. Such revelations can be enormously important, leading to your altering or modifying your original research design.

As examples, in a single-case study, what was thought to be a critical or unusual case might have turned out not to be so, after initial data collection had started; ditto a multiple-case study, where what was thought to be parallel cases for literal replication turn out not to be so. With these revelations, you have every right to conclude that your initial design needs to be modified. However, you should undertake any alterations only given a serious caution. The caution is to understand precisely the nature of the alteration: Are you merely selecting different cases, or are you also changing your original theoretical concerns and objectives? The point is that the needed adaptiveness should not lessen the rigor with which case study procedures are followed.

Mixed Methods Designs: Mixing Case Studies with Other Methods?

Researchers have given increasing attention to *mixed methods research*—a "class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a *single* study" (Johnson & Onwuegbuzie, 2004, p. 17, emphasis added). Confinement to a single study forces the methods being mixed into an integrated mode. The mode differs from the conventional situation whereby different methods are used in *separate* studies that may later be synthesized.

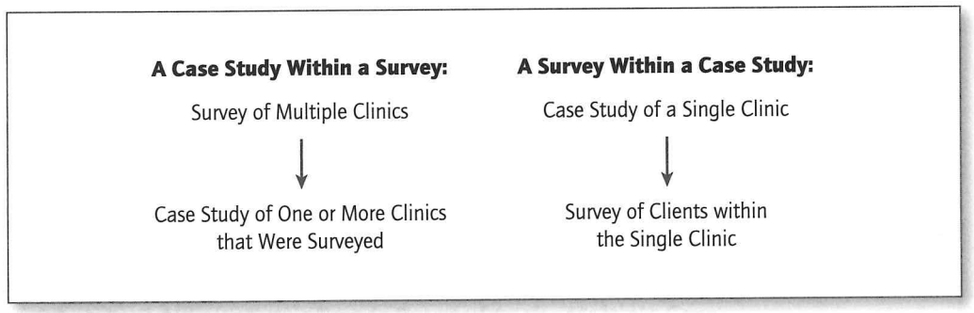
Mixed methods research forces the methods to share the same research questions, to collect complementary data, and to conduct counterpart analyses (e.g., Yin, 2006b)—in short, to follow a mixed methods design. As such, mixed methods research can permit

researchers to address more complicated research questions and collect a richer and stronger array of evidence than can be accomplished by any single method alone. Depending upon the nature of your research questions and your ability to use different methods, mixed methods research opens a class of research designs that deserve your attention.

The earlier discussion of embedded case study designs in fact points to the fact that certain kinds of case studies already represent a form of mixed methods research: Embedded case studies may rely on holistic data collection strategies for studying the main case and then call upon surveys or other quantitative techniques to collect data about the embedded unit(s) of analysis. In this situation, other research methods are embedded within case study research.

The opposite relationship also can occur. Your case study may be part of a larger, mixed methods study. The main investigation may rely on a survey or other quantitative techniques, and your case study may help to investigate the conditions within one of the entities being surveyed. The contrasting relationships (survey within case or case within survey) are illustrated in Figure 2.6 (also see Appendix B for further discussion of these mixtures in relation to evaluation studies).

Figure 2.6 Mixed Methods: Two Nested Arrangements



At the same time, mixed methods research need not include the use of case study research at all. For instance, a clinical study could be combined with historical work that embraces the quantitative analysis of archival records, such as newspapers and other file material. Going even further, mixed methods research need not be limited to combinations of quantitative and qualitative methods. For instance, a study could employ a mix of two quantitative methods: a survey to describe certain conditions, complemented by an experiment that tries to manipulate some of those conditions (e.g., Berends & Garet, 2002).

By definition, studies using mixed methods research are more difficult to execute than studies limited to single methods. However, mixed methods research can enable you to address broader or more complicated research questions than case studies alone. As a result, mixing case study research with other methods should be among the possibilities meriting your consideration.

NOTES TO CHAPTER 2 △

1. Figure 2.2 focuses only on the formal research design process, not on data collection activities. For all three types of research (survey, case study, and experiment), data collection techniques might be depicted as the level below Level One in the figure. For example, for case study research, this might include using multiple sources of evidence, as described further in Chapter 4. Similar data collection techniques can be described for surveys or experiments—for example, questionnaire design for surveys or stimulus presentation strategies for experiments.

2. Whether experiments also need to address statistical generalizations has been the topic of sharp debate in psychology. According to the statistical argument, the human subjects in an experiment should be considered a population sample, with the experimental results therefore limited to the universe of the same population. The debate began over the excessive use of college sophomores in behavioral research (e.g., Cooper, McCord, & Socha, 2011; Gordon, Slade, & Schmitt, 1986; McNemar, 1946; Peterson, 2001; Sears, 1986) and has since extended to an awareness that the subjects in most behavioral research have been white males from industrialized countries (Henrich, Heine, & Norenzayan, 2010), even though the experimental findings are intended to apply as "the norm for all human beings" (Prescott, 2002, p. 38).

3. One of the anonymous reviewers of the third edition of this book pointed out that construct validity also has to do with whether interviewees understand what is being asked of them.

4. For other suggested guidelines for reviewers of case study proposals or manuscripts, see Yin (1999).

5. Strictly quantitative studies that select cases with known outcomes follow the same design and have alternatively been called "case-control," "retrospective," or "case referent" studies (see Rosenbaum, 2002, p. 7).

Tutorial 2.1: More on Defining "Analytic Generalization"

An analytic generalization consists of a carefully posed theoretical statement, theory, or theoretical proposition. The generalization can take the form of a lesson learned, working hypothesis, or other principle that is believed to be applicable to other situations (not just other "like cases"). Thus, the preferred analytic generalization is posed at a conceptual level higher than that of the specific case (presumably, this higher level was needed to justify the importance of studying the chosen case in the first place).

Though not using the same terminology, other prominent works have devoted attention to analytic generalization, also distinguishing it from statistical generalization: (1) Mitchell's (1983) discussion of *logical inference* and *statistical inference;* (2) Bromley's (1986) discussion of *case inference* compared with *statistical inference* (pp. 290–291); and Donmoyer's (1990) *schema.* A fourth work, by Burawoy (1991), covers the *extended case method*—his way of describing how a generalization "extends" a narrow case to some broader significance (pp. 271–280).

The more difficult and contrary position—that the studied case should be construed as an instance, example, or sample of some larger group of cases—undesirably returns to statistical generalization (the relationship between a sample and its population—e.g., Gomm, Hammersley, & Foster, 2000, pp. 99–103). That position dwells on the fact that a "case" seems to be an instance or example of other "like cases." However, such a claim is inappropriate when thinking about analytic generalization, where the findings from a case study can have implications going well beyond the same kind of case and extend to a whole host of other unlike situations (see BOX 7, p. 43, in the main text for three examples). Moreover, unless a case study has included a large number of cases—typically dozens or scores if not hundreds of cases (see Tutorial 5-3)—the study will face an uphill battle by invoking the sample to population analogy and its concomitant need to employ statistical analyses to assess the strength of any relationship.

Small (2009) provides two excellent examples and an insightful discussion of analytic generalization, also citing the same key works as referenced above. To him, the preferred logic represents "a different perspective and language of inquiry" (p. 18). He further notes the importance of starting with a substantive proposition (e.g., a causal relationship) rather than a numeric one (e.g., the representativeness of a case) to make analytic generalizations work.

Briefly Annotated References for Tutorial 2.1

Bromley, D. B. (1986). *The case-study method in psychology and related disciplines*. Chichester, England: Wiley. Provides comprehensive guidance on case study research in psychology.

Burawoy, M. (1991). The extended case method. In M. Burawoy, A. Burton, A. A. Ferguson, K. J. Fox, J. Gamson, N. Gartrell, et al. (Eds.), *Ethnography unbound: Power and resistance in the modern metropolis* (pp. 271–287). Berkeley: University of California Press. Presents the extended case method for analyzing participant-observation data.

Donmoyer, R. (1990). Generalizability and the single-case study. In E. W. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate* (pp. 175–200). New York: Teachers College Press. Offers a way of generalizing from single studies, not based on sampling and statistical significance.

Gomm, R., Hammersley, M., & Foster, P. (2000). Case study and generalization. In R. Gomm, M. Hammersley, & P. Foster (Eds.), *Case study method* (pp. 98–115). London: Sage. Highlights use of the case study method for generalizing, rather than merely studying a case for its own sake.

Mitchell, J. C. (1983). Case and situation analysis. *Sociological Review, 31,* 187–211. Emphasizes case study research as a method for preserving the unitary character of the social object being studied and discusses the challenge of generalizing from the case(s).

Small, M. L. (2009). "How many cases do I need?" On science and the logic of case selection in field-based research. *Ethnography, 10,* 5–38. Poses a thoughtful article on key issues in designing field-based research, including the challenge of generalizing from field situations.