2

Designing Case Studies

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

GENERAL APPROACH TO DESIGNING CASE STUDIES

In identifying the method for your research project, Chapter 1 has shown when you might choose to use the case study method, as opposed to other methods. The next task 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 is a difficult part of doing case studies. Unlike other research methods, a comprehensive "catalog" of research designs for case studies has yet to be developed. There are no textbooks, like those in the biological and psychological sciences, covering such design considerations as the assignment of subjects to different "groups," the selection of different stimuli or experimental conditions, or the identification of various response measures (see Cochran & Cox, 1957; Fisher, 1935, cited in Cochran & Cox, 1957; Sidowski, 1966). In a laboratory experiment, each of these choices reflects an important logical connection to the issues being studied. Similarly, there are not even textbooks like the well-known volumes by Campbell and Stanley (1966) or by Cook and Campbell (1979) that summarize the various research designs for quasi-experimental situations. Nor have there emerged any common designs—for example, "panel" studies—such as those recognized in doing survey research (see L. 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 experiments. For the longest time, scholars incorrectly thought that the case study was but one type of quasi-experimental design (the "one-shot post-test-only" design). This misperception has finally been corrected, with the following statement appearing in a revision on quasi-experimental designs (Cook & Campbell, 1979): "Certainly the case study as normally practiced should not be demeaned by identification with the one-group post-test-only design" (p. 96). In other

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

You need sufficient access to the potential data, whether to interview people, review documents or records, or make observations in the "field." 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?

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, the case study is a separate research method that has its own research designs.

Unfortunately, case study research designs have not been codified. The following chapter therefore expands on the new methodological ground broken by 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 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 answered only if you collect information directly from the other organizations and not merely from the one you started with. If you complete your study by examining only one organization, you cannot draw unbiased conclusions about interorganizational partnerships. This is a flaw in your research design, not in your work plan. The outcome could have been avoided if you had developed an appropriate research design in the first place.

Components of Research Designs

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

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

Study questions. This first component has already been described in Chapter 1, 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. The case study method 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 coming 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 provide support for your potential questions or even suggest ways of sharpening them.

EXERCISE 2.1 Defining the Boundaries of a Case Study

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. How does the naming of these questions or propositions clarify the boundaries of your case study 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?

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 answering, led you to the case study as the appropriate method in the first place. Nevertheless, these "how" and "why" questions do not point to what you should study.

Only if you are forced to state some propositions will you move in the right direction. For instance, you might think that organizations collaborate because they derive mutual benefits. This proposition, 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 (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. Consider the analogy in BOX 4 for exploratory case studies. Can you imagine how you would ask for support from Queen Isabella to do your exploratory study?

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

When Christopher Columbus went to Queen Isabella to ask for support for his "exploration" of the New World, he had to have some reasons for asking for three ships (Why not one? Why not five?), and he had some rationale for going westward (Why not south? Why not south and then east?). He also had some (mistaken) criteria for recognizing the Indies when he actually encountered 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. This third component is related to the fundamental problem of defining what the "case" is—a problem that has plagued many investigators at the outset of case studies (e.g., Ragin & Becker, 1992). For instance, in the classic case study, a "case" may be an individual. Jennifer Platt (1992) has noted how the early case studies 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 involve 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 decisions, programs, the implementation process, and organizational change. 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.

As a general guide, your tentative definition of the unit of analysis (which is the same as the definition of the "case") is related to the way you have defined your initial research questions. 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 (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. Using Drucker's work or some similar theoretical framework, the unit of analysis (or "case") for your case study 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 slightly different research design and data collection strategy.

Selection of the appropriate unit of analysis will start to occur when you accurately specify your primary research questions. If your questions do not lead to the favoring of one unit of analysis over another, your questions are probably either too vague or too numerous—and you may have trouble 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 flexibility 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 5A). 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, 1943/1955; 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 implications). BOX 5B, however, presents a good example of how units of analyses can be defined in a more discriminating manner—in the field of world trade.

BOX 5
Defining the Unit of Analysis

5A. What Is the Unit of Analysis?

The Soul of a New Machine (1981) was a Pulitzer Prize–winning book by Tracy Kidder. The book, also a best seller, is about the development of a new minicomputer, produced by Data General Corporation, intended to compete with one produced by a direct competitor, Digital Equipment Corporation (also see BOX 28, Chapter 5, p. 142).

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

The book is an excellent example of a case study. However, the book also illustrates a fundamental problem in doing case studies—that of defining the unit of analysis. Is the "case" 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.

5B. 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 35, Chapter 5, p. 161). 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, about the Korean firm Samsung, is a case study of the critical policies that make the firm competitive. Understanding Korean economic development is part of the context, and the case study also contains an embedded unit—Samsung's development of the microwave oven as an illustrative product. The other case, about the development of an Apple computer factory in Singapore, is in fact a case study of Singapore's critical policies that make the country competitive. The Apple computer factory experience—an embedded unit of analysis—is actually an illustrative example of how the national policies affected foreign investments.

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

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

Once the general definition of the case has been established, other clarifications in the unit of analysis become important. If the unit of analysis is a small group, for instance, the persons to be included within the group (the immediate topic of the case study) must be distinguished from those who are outside it (the context for the case study). Similarly, if the case is about 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 specific time boundaries to define the beginning and end of the case (e.g., whether to include the entire or only some part of the life cycle of the entity that is to be the case). Answering all of these types of questions 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").

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, not an abstraction such as a topic, an argument, or even a hypothesis. These abstractions, absent the identification of specific examples or cases, would rightfully serve as the subjects of research studies using other kinds of methods but not case studies. To justify using the case study method, you need to go one step further: You need to define a specific, real-life "case" to represent 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 setting of boundaries, or an experiment in which young children do tasks next to each other to determine the distracting effects, if any, of their neighbors. 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

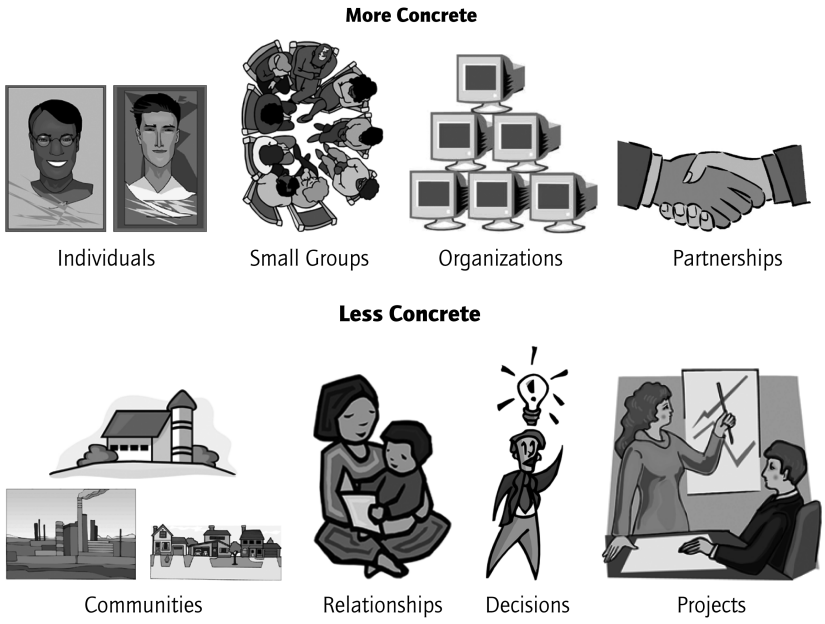


Figure 2.1 Illustrative Case Study Topics

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.

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?

Linking data to propositions and criteria for interpreting the findings. The fourth and fifth components have been increasingly better developed in doing case studies. These components foreshadow the data analysis steps in case study research. Because the analytic techniques and choices are covered in

detail in Chapter 5, your main concern during the design phase is to be aware of the main 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.

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 calculate 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. Noting this strong likelihood during the design phase would call your attention to the need to be sure you had sufficient procedures to collect time markers as part of your data collection plans.

If you have had limited experience in conducting empirical studies, you will not easily identify the likely analytic technique(s) or anticipate the needed data to use the techniques to their full advantage. More experienced researchers will note how often 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 any of these situations, the better off you will be.

Criteria for interpreting a study's findings. Statistical analyses offer some explicit criteria for such interpretations. For instance, by convention, social science considers a p level of less than .05 to demonstrate that observed differences were "statistically significant." However, much case study analysis will not rely on the use of statistics and therefore calls attention to other ways of thinking about such criteria.

A major and important alternative strategy is to identify and address rival explanations for your findings. Again, Chapter 5 discusses this strategy and how it works more fully. At the design stage of your work, the challenge is to anticipate and enumerate the important rivals, so you will include information 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. Although the current state of the art does not provide detailed guidance on the last two, the complete research design should indicate what data are to be collected—as

indicated by a study's questions, its propositions, and its units of analysis. The design also should tell you what is to be done after the data have been collected—as indicated by the logic linking the data to the propositions and the criteria for interpreting the findings.

The Role of Theory in Design Work

Covering these preceding five components of research designs will effectively force you to begin constructing a preliminary theory related to your topic of study. This role of theory development, prior to the conduct of any data collection, is one point of difference between case studies and related methods such as ethnography (Lincoln & Guba, 1985; Van Maanen, 1988) and "grounded theory" (Corbin & Strauss, 2007). Typically, these related methods deliberately avoid specifying any theoretical propositions at the outset of an inquiry. As a result, students confusing these methods with case studies wrongly think that, by having selected the case study method, they can proceed quickly into the data collection phase of their work, and they may have been encouraged to make their "field contacts" as quickly as possible. No guidance could be more misleading. Among other considerations, the relevant field contacts depend upon an understanding—or theory—of what is being studied.

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

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

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

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

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

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

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

You can see that as these two initial ingredients are elaborated, the stated ideas will increasingly cover the questions, propositions, units of analysis, logic connecting data to propositions, and criteria for interpreting the findings—that is, the five components of the needed research design. In this sense, the complete research design embodies a "theory" of what is being studied.

This theory should by no means be considered with the formality of grand theory in social science, nor are you being asked to be a masterful theoretician. Rather, the simple goal is to have a sufficient blueprint for your study, and this requires theoretical propositions, usefully noted by Sutton and Staw (1995) as "a [hypothetical] story about why acts, events, structure, and thoughts occur" (p. 378). Then, the complete research design will provide surprisingly strong guidance in determining what data to collect and the strategies for analyzing the data. For this reason, theory development prior to the collection of any case study data is an essential step in doing case studies. As noted for nonexperimental studies more generally, a more elaborate theory desirably points to a more complex pattern of expected results (P. R. 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" 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 4, 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 (2003, chap. 1) shows how theory was used in exploratory, descriptive, and explanatory situations by discussing five actual case studies.

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

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

- 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 behavior, 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 studies is the evaluation of publicly supported programs, such as federal, state, or local programs. In this situation, the development of a theory of how a program is supposed to work is essential to the design of the evaluation. 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.

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

Let us first take the more commonly recognized way of generalizing—*statistical* generalization—although it is the less relevant one for doing case studies. In statistical generalization, an inference is made about a population (or universe) on the basis of empirical data collected about a sample from that universe. This is shown as a Level One inference in Figure 2.2.[1] This method of generalizing is commonly recognized because research investigators have ready access to quantitative formulas for determining the confidence with which generalizations can be made, depending mostly upon the size and internal variation within the universe and sample. Moreover, this is the most common way of generalizing when doing surveys (e.g., Fowler, 1988; Lavrakas, 1987) or analyzing archival data.

A fatal flaw in doing case studies is to conceive of statistical generalization as the method of generalizing the results of your case study. This is because your cases are not "sampling units" and should not be chosen for this reason. Rather, individual case studies are to be selected as a laboratory investigator selects the topic of a new experiment. Multiple cases, in this sense, resemble multiple experiments. Under these circumstances, the mode of generalization is *analytic* generalization, in which a previously developed theory is used as a template with which to compare the empirical results of the case study.[2] If two or more cases are shown to support the same theory, replication may be

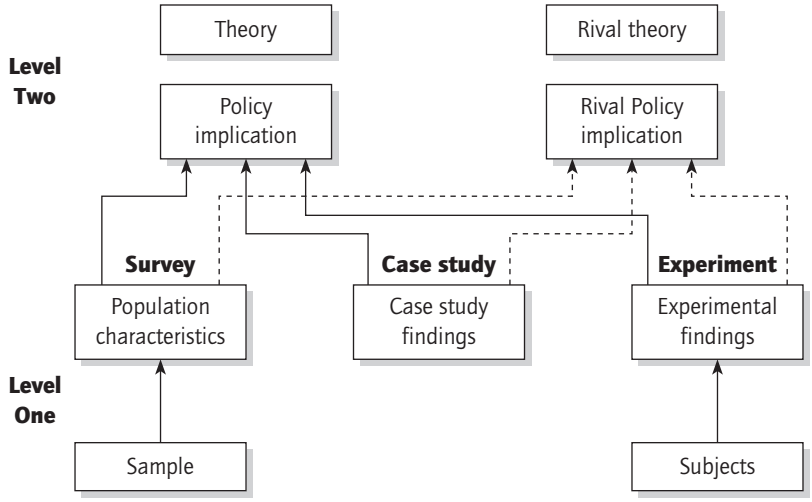


Figure 2.2 Making Inferences: Two Levels

claimed. The empirical results may be considered yet more potent if two or more cases support the same theory but do not support an equally plausible, *rival* theory. Graphically, this type of generalization is shown as a Level Two inference in Figure 2.2.

Analytic generalization can be used whether your case study involves one or several cases, which shall be later referenced as single-case or multiple-case studies. Furthermore, the logic of replication and the distinction between statistical and analytic generalization will be covered in greater detail in the discussion of multiple-case study designs. The main point at this juncture is that you should try to aim toward analytic generalization in doing case studies, and you should avoid thinking in such confusing terms as "the sample of cases" or the "small sample size of cases," as if a single-case study were like a single respondent in a survey or a single subject in an experiment. In other words, in terms of Figure 2.2, you should aim for Level Two inferences when doing case studies.

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

Summary. This subsection has suggested that a complete research design, covering the four components described earlier, in fact requires the development

of a theoretical framework for the case study that is to be conducted. Rather than resisting such a requirement, a good case study investigator should make the effort to develop this theoretical framework, no matter whether the study is to be explanatory, descriptive, or exploratory. The use of theory, in doing case studies, is an immense aid in defining the appropriate research design and data collection. The same theoretical orientation also becomes the main vehicle for generalizing the results of the case study.

CRITERIA FOR JUDGING THE QUALITY OF RESEARCH DESIGNS

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

Four tests, however, have been commonly used to establish the quality of any empirical social research. Because case studies are one form of such research, the four tests also are relevant to case studies. An important innovation of this book is the identification of several tactics for dealing with these four tests when doing case studies. Figure 2.3 lists the four widely used tests and the recommended case study tactics, as well as a cross-reference to the phase of research when the tactic is to be used. (Each tactic is 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 (see L. Kidder & Judd, 1986, pp. 26–29):

- *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—such as the data collection procedures—can be repeated, with the same results

Each item on this list deserves explicit attention. For case studies, an important revelation is that the several tactics to be used in dealing with these tests

TESTS	Case Study Tactic	Phase of research in which tactic occurs
Construct validity	♦ use multiple sources of evidence	data collection
	♦ establish chain of evidence	data collection
	♦ have key informants review draft case study report	composition
Internal validity	♦ do pattern matching	data analysis
	♦ do explanation building	data analysis
	♦ address rival explanations	data analysis
	♦ use logic models	data analysis
External validity	♦ use theory in single-case studies	research design
	♦ use replication logic in multiple-case studies	research design
Reliability	♦ use case study protocol	data collection
	♦ develop case study database	data collection

Figure 2.3 Case Study Tactics for Four Design Tests

should be applied throughout the subsequent conduct of the case study, not just at its beginning. Thus, the "design work" for 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 investigator fails to develop a sufficiently operational set of measures and that "subjective" judgments 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 an investigator'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 Figure 2.3 shows, 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 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. This is especially true in doing case studies. As one set of suggestions, Figure 2.3 shows that the analytic tactic of *pattern matching,* described further in Chapter 5, is one way of addressing internal validity. Three other analytic tactics, *explanation building, addressing rival explanations,* and *using logic models,* also are described in Chapter 5.

External Validity

The third test deals with the problem of knowing whether a study's findings are generalizable beyond the immediate case study. In the simplest example, if a study of neighborhood change focused on one neighborhood, are the results applicable to another neighborhood? The external validity problem has been a major barrier in doing case studies. Critics typically state that single cases offer a poor basis for generalizing. However, such critics are implicitly contrasting the situation to survey research, in which a sample is intended to generalize to a larger universe. *This analogy to samples and universes is incorrect when dealing with case studies.* Survey research relies on *statistical* generalization, whereas case studies (as with experiments) rely on *analytic* generalization. In analytical generalization, the investigator is striving to generalize a particular set of results to some broader theory (see three examples in BOX 6).

For example, the theory of neighborhood change that led to a case study in the first place is the same theory that will help to identify the other cases to which the results are generalizable. If a study had focused on population transition in an urban neighborhood (e.g., Flippen, 2001), the procedure for selecting a neighborhood for study would have begun with identifying a neighborhood within which the hypothesized transitions were occurring. Theories about transition would then be the domain to which the results could later be generalized.

BOX 6
How Case Studies Can Be Generalized to Theory: Three Examples

6A. The Origins of Social Class Theory

The first example is about the uncovering and labeling of a social class structure based on a case study of a typical 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 social differences among "upper," "upper-middle," "middle-middle," "upper-lower," and "lower" classes.

6B. Contributions to Urban Planning Theory

The second 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. However, the chapter topics, rather than reflecting the single experiences of New York, cover broader theoretical issues in urban planning, such as the role of sidewalks, the role of neighborhood parks, the need for primary mixed uses, the need for small blocks, and the processes of slumming and unslumming. In the aggregate, these issues in fact represent Jacobs's building of a theory of urban planning.

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

6C. A More Contemporary Example

A third example covers a 5-year ethnographic study of a single neighborhood at the edge of Chicago (Carr, 2003). The study shows how the neighborhood successfully thwarted undesirable youth-related crime. The experience, in the author's view, challenged existing theories claiming that strong social ties are crucial to effective neighborhood control. Instead, the author offers newer theories of informal social control that he believes may be especially pertinent to youth crime prevention in contemporary suburban neighborhoods.

The generalization is not automatic, however. A theory must be tested by replicating the findings in a second or even a third neighborhood, where the theory has specified that the same results should occur. Once such direct replications have been made, the results might be accepted as providing strong support for the theory, even though further replications had not been performed. This *replication logic* is the same that underlies the use of experiments (and allows scientists to cumulate knowledge across experiments). The logic will be discussed further in this chapter in the section on multiple-case designs.

Reliability

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

One prerequisite for allowing this other investigator to repeat an earlier case study is the need to document the procedures followed in the earlier case. Without such documentation, you could not even repeat your own work (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] Figure 2.3 indicates two specific tactics to overcome 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 always 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 arrive at the same results.

Summary

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

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 DESIGNS

These general characteristics of research designs serve as a background for considering the specific designs for case studies. Four types of designs will be discussed, based on a 2 × 2 matrix (see Figure 2.4). The matrix 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)

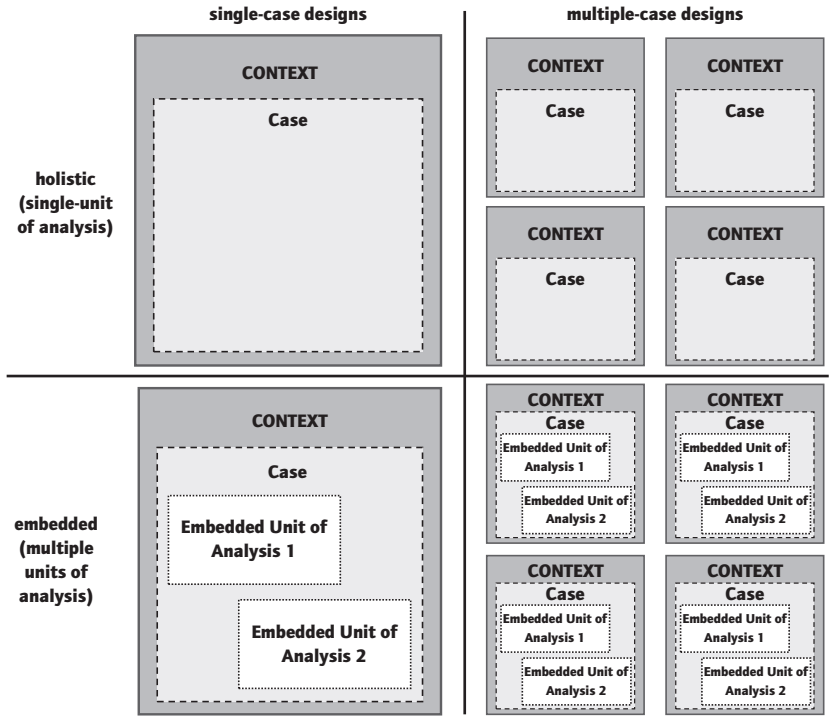


Figure 2.4 Basic Types of Designs for Case Studies
SOURCE: COSMOS Corporation.

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)?

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

One rationale for a single case is when it represents the *critical case* in testing a well-formulated theory (again, note the analogy to the critical *experiment*). The theory has specified a clear set of propositions as well as the circumstances within which the propositions are believed to be true. A single case, meeting all of the conditions for testing the theory, can confirm, challenge, or extend the theory. The single case can then be used to determine whether a theory's propositions are correct or whether some alternative set of explanations might be more relevant. In this manner, like Graham Allison's comparison of three theories and the Cuban missile crisis (described in Chapter 1, BOX 2), the single case can represent a significant contribution to knowledge and theory building. Such a study can even help to refocus future investigations in an entire field. (See BOX 7 for another example, in the field of organizational innovation.)

A second rationale for a single case is where the case represents an *extreme* case or a *unique* case. Either of these situations commonly occurs in clinical psychology, where a specific injury or disorder may be so rare that any single case is worth documenting and analyzing. For instance, one rare clinical syndrome is the inability of certain clinical patients to recognize familiar faces. Given visual cues alone, such patients are unable to recognize loved ones, friends, pictures of famous people, or (in some cases) their own image in a mirror. This syndrome appears to be due to some physical injury to the brain. Yet the syndrome occurs so rarely that scientists have been unable to establish any common patterns (Yin, 1970, 1978). In such circumstances, the single-case study is an appropriate research design whenever a new person with this syndrome—known as prosopagnosia—is encountered. The case study would document the person's abilities and disabilities, determine the precise nature of the face recognition deficit, but also ascertain whether related disorders exist.

BOX 7
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 19B, Chapter 3, p. 110).

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.

Conversely, a third rationale for a single case is the *representative* or *typical* case. Here, the objective is to capture the circumstances and conditions of an everyday or commonplace situation (see BOX 8; also see BOX 14, p. 75). The case study may represent a typical "project" among many different projects, a manufacturing firm believed to be typical of many other manufacturing firms in the same industry, a typical urban neighborhood, or a representative school, as examples. The lessons learned from these cases are assumed to be informative about the experiences of the average person or institution.

A fourth rationale for a single-case study is the *revelatory* case. This situation exists when an investigator has an opportunity to observe and analyze a phenomenon previously inaccessible to social science inquiry, such as Whyte's

BOX 8
The Average Case as a Single-Case Study

A famous community case study in sociology, *Middletown,* is about an average American city. The investigators, Robert and Helen Lynd (1929), deliberately chose to study a small town in middle America during the early 20th century (also see BOX 14, p. 75). Their purpose was to show how the transition from an agricultural to an industrial economy occurred in the average town—and thereby to provide a case study about a significant development in all of American history.

(1943/1955) *Street Corner Society,* previously described in Chapter 1, BOX 2A. Another example is Elliot Liebow's (1967) famous case study of unemployed men, *Tally's Corner* (see BOX 9). Liebow had the opportunity to meet the men in an African American neighborhood in Washington, D.C. and to learn about their everyday lives. His observations of and insights into the problems of unemployment formed a significant case study, because few social scientists had previously had the opportunity to investigate these problems, even though the problems were common across the country. When other investigators 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 rather than a multiple-case design is that the investigator has access to a situation previously inaccessible to scientific observation. The case study is therefore worth conducting because the descriptive information alone will be revelatory.

Such was the situation in Elliot Liebow's (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 coping behavior, and in particular their sensitivity to unemployment and failure. The book provided insights into a subculture that has prevailed in many U.S. cities for a long period of time, but one that had been only obscurely understood. The single case showed how investigations of such topics could be done, thus stimulating much further research and eventually the development of 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 change over time, and the desired time intervals would presumably reflect the anticipated stages at which the changes should reveal themselves.

These five serve as major reasons for conducting a single-case study. There are other situations in which the single-case study may be used as a pilot case that is the first of a multiple-case study. However, in these latter instances, the single-case study cannot be regarded as a complete study on its own.

Whatever the rationale for doing single-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 all of these major concerns have been covered.

Holistic versus embedded case studies. The same single-case study may involve *more than one unit of analysis.* 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 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. 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 an investigator 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.

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). 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 (BOX 10 Continued)

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

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

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 the case study approach, 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 cases are a common design for doing case studies, 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) a rare or unique circumstance, or (c) a representative or typical 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 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 the case study method.

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 rare 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 mistaken analogy in the past, which incorrectly considered multiple cases to be similar to the multiple respondents in a survey (or to the multiple subjects within

an experiment)—that is, to follow a "sampling" 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 scientists deal with conflicting experimental findings.

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

For example, one might consider the initial proposition that an increase in using a new technology in school districts will occur when the technology is used for both administrative and instructional applications, but not either alone. To pursue this proposition in a multiple-case study design, 3 or 4 cases might be selected in which both types of applications are present, to determine whether, in fact, technology use 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 instructional 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 HIV/AIDS prevention, university administration, and the transformation of business firms, in the companion text (Yin, 2003, chaps. 8–10).

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

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

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 an investigator wishes to determine the prevalence or frequency of a particular phenomenon.

Any application of this sampling logic to case studies 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 number of cases—too large to allow any statistical consideration of the relevant variables.

Third, if a sampling logic had to be applied to all types of research, many important topics could not be empirically investigated, such as 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 even may require you to reconsider one or more of the study's original theoretical propositions. At this point,

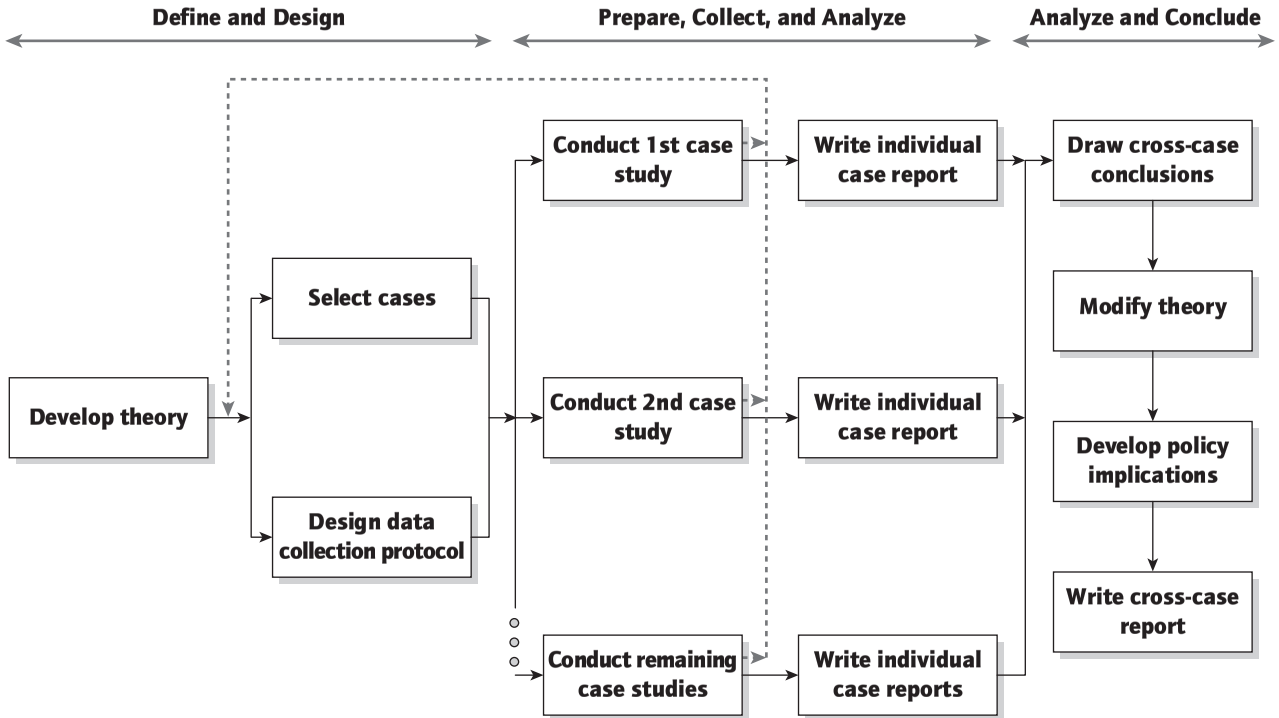


Figure 2.5 Case Study Method

SOURCE: COSMOS Corporation.

"redesign" should take place before proceeding further. Such redesign might involve the selection of alternative cases or changes in the case study (i.e., data collection) 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 (i.e., the original theoretical propositions).

Overall, Figure 2.5 depicts a very 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 sample size also are irrelevant. Instead, you should think of this decision as a reflection of the number of case replications—both literal and theoretical—that you need or would like to have in your study.

For the number of literal replications, an appropriate analogy from statistics is the selection of the criterion for establishing the sample size desired to detect an "effect." Designating a "$p < .05$" or "$p < .01$" likelihood of detection as part of a power analysis is not based on any formula but is a matter of discretionary, judgmental choice. Analogously, designating the number of replications depends upon the certainty you want to have about your multiple-case results (as with the higher criterion for establishing the likelihood of detection, the greater certainty lies with the larger number of cases). 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 high degree of certainty, you may press for five, six, or more replications.

For the number of theoretical replications, the important consideration is related to your sense of the 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 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.

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 questions, 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 good and 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 cases: Each individual case may still be holistic or embedded. In other words, a multiple-case study may consist of multiple holistic cases (see Figure 2.4, Type 3) or of multiple embedded cases (see Figure 2.4, Type 4).

The difference between these two 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 schools. Each 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 school, or case. These data may be highly quantitative, focusing on the attitudes and behavior of individual students, and the data will be used along with archival information to interpret the success and operations at the given 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 a single-case study, in which all nine schools and their students have now become part of some larger, main unit of analysis. Such a new case study would then require a complete redefinition of the main unit of analysis, with extensive revisions to the original theories and propositions of interest also a likely need.

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

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

The 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.

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 basic research questions to be answered? What methods were used to draw conclusions, based on the evidence? Is the design a single- or multiple-case design? Is it holistic or does it have embedded units of analysis?

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, three pieces of advice may be offered.

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. Even if you can do 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 vastly strengthening your findings compared to those from a single case 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 Designs or Flexible 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 design.

As examples, in a single-case study, what was thought to be a critical or unique 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 flexibility 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 investigators 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 consideration.

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. The embedded case studies rely on more holistic data collection strategies for studying the main case but then call upon surveys or other more quantitative techniques to collect data about the embedded unit(s) of analysis. In this situation, other research methods are embedded within your case study.

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.

At the same time, mixed methods research need not include the use of the case study strategy at all. For instance, much historical work embraces the quantitative analysis of archival records, such as newspapers and other file material. And, in an even broader sense, mixed methods research need not be limited to combinations of quantitative and qualitative methods. For instance, a study could employ a survey to describe certain conditions, complemented by an experiment that tried to manipulate some of those conditions (e.g., Berends & Garet, 2002).

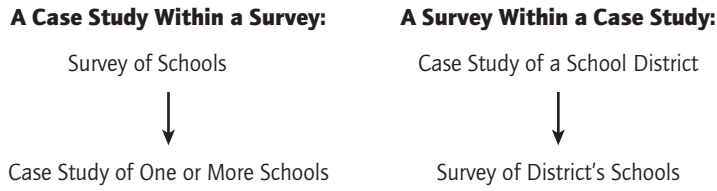


Figure 2.6 Mixed Methods: Two Nested Arrangements

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 studies with other methods should be among the possibilities meriting your consideration.

NOTES

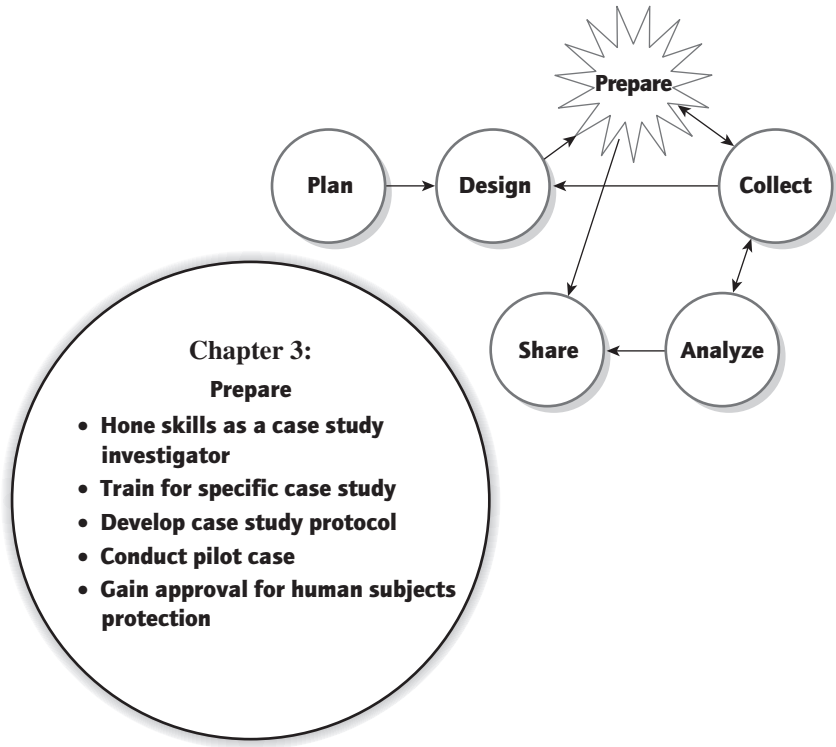
1. Figure 2.2 focuses only on the formal research design process, not on data collection activities. For all three types of research (survey, case study, and experiment), data collection techniques might be depicted as the level below Level One in the figure. For example, for case studies, this might 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. See Gomm, Hammersley, and Foster (2000) for more explanation of *analytic* generalization, though their work uses different labels for the same concept.
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 P. R. Rosenbaum, 2002, p. 7).

REFERENCE TO EXPANDED CASE STUDY MATERIALS FOR CHAPTER 2

For selected case studies cited in the text of this chapter, two anthologies contain either a more extensive excerpt or the full case study. The table on the next page crosswalks the reference in this book to the location of the excerpt or full rendition.

Chapter 2 *Chapter Topic and Page Number*	*Topics of Illustrative Case Studies*	*Reference to Lengthier Material*
General Approach to Designing Case Studies		
BOX 4, p. 2-6	Exploratory study	None
BOX 5A, p. 2-9	Computers and technology	None
BOX 5B, p. 2-9	Business and industry	CSA-6
p. 2-15 text	Five different case studies	ACSR-1
Criteria for Judging the Quality of Research Designs		
BOX 6A, p. 2-23	Cities and towns	CSA-4
BOX 6B, p. 2-23	Urban planning	None
BOX 6C, p. 2-23	Neighborhoods	None
Case Study Designs		
BOX 7, p. 2-27	Schools	CSA-9
BOX 8, p. 2-28	Cities and towns	CSA-3
BOX 9, p. 2-28	Neighborhoods	None
BOX 10, p. 2-29	Business and industry	CSA-10
BOX 11, p. 2-35	Government agencies	None
p. 2-35 text	Health (HIV/AIDS) care	ACSR-8
p. 2-35 text	University administration	ACSR-9
p. 2-35 text	Business and industry	ASCR-10
Modest Advice in Selecting Case Study Designs		
BOX 12A, p. 2-41	Community organizations	None
BOX 12B, p. 2-41	Schools	None

NOTE: CSA = *Case Study Anthology* (Yin, 2004). ACSR = *Applications of Case Study Research* (Yin, 2003). The number denotes the chapter number in the book.



ABSTRACT

Preparing to do a case study starts with the prior skills of the investigator and covers the preparation and training for the specific case study (including procedures for protecting human subjects), the development of a case study protocol, the screening of candidate cases to be part of the case study, and the conduct of a pilot case study.

With regard to prior skills, many people incorrectly believe they are sufficiently skilled to do case studies because they think the method is easy to use. In fact, case study research is among the hardest types of research to do because of the absence of routine procedures. Case study investigators therefore need to feel comfortable in addressing procedural uncertainties during the course of a study. Other desirable traits include the ability to ask good questions, "listen," be adaptive and flexible, have a firm grasp of the issues being studied, and know how to avoid bias.

An investigator can prepare to do a high-quality case study through intensive training. A case study protocol should be developed and refined. These procedures are especially desirable if the research is based on a multiple-case design or involves multiple investigators, or both.