



large number of near-synonyms – single unit, single subject, single case, N = 1, case-based, case-control, case history, case method, case record, case work, within-case, clinical research, and so forth.<sup>9</sup> As a result of this profusion of terms and meanings, proponents and opponents of the case study marshal a wide range of arguments but do not seem any closer to agreement than when this debate was first broached several decades ago. Jennifer Platt notes that “much case study theorizing has been conceptually confused, because too many different themes have been packed into the idea ‘case study.’”<sup>10</sup>

How, then, should the case study be understood? The first six options enumerated above (a–f) seem inappropriate as general definitions of the topic, since each implies a substantial shift in meaning relative to established usage. One cannot substitute case study for *qualitative, ethnographic, process-tracing, holistic, naturalistic, diffuse, or triangulation* without feeling that something has been lost in translation. These terms are perhaps better understood as describing certain kinds of case studies, not the topic at large. A seventh option, (g), equates the case study with the study of a single observation, the N = 1 research design. This is logically impossible, as I will argue. The eighth option, (h), centering on *phenomenon, instance, or example* as the key term, is correct as far as it goes but also ambiguous. Imagine asking someone, “What is your instance?” or “What is your phenomenon?” A case study presupposes a relatively bounded phenomenon, an implication that none of these terms captures.

Can this concept be reconstructed in a clearer, more productive fashion? I begin this chapter by stipulating a series of definitions. I then present a typology of research designs, understood according to the patterns of spatial and temporal evidence that they draw upon. A final section addresses a central definitional question, namely, whether case studies should be understood as exclusively “small-N” analyses.

<sup>9</sup> Davidson and Costello (1969); Franklin, Allison, and Gorman (1997); Hersen and Barlow (1976); Kazdin (1982); Kraotchwill (1978).

<sup>10</sup> Platt (1992: 48). Elsewhere in this perceptive article, Platt (1992: 37) comments: “the diversity of the themes which have been associated with the term, and the vagueness of some of the discussion, causes some difficulty... In practice, ‘case study method’ in its heyday [in the interwar years] seems to have meant some permutation of the following components: life history data collected by any means, personal documents, unstructured interview data of any kind, the close study of one or a small number of cases whether or not an attempt was made to generalize from them, any attempt at holistic study, and non-quantitative data analysis. These components have neither a necessary logical nor a regular empirical connection with each other.”

## Definitions

For purposes of methodological discussion, it is essential to develop a vocabulary that is consistent and clear. In arriving at definitions for key terms, I rely on ordinary usage (within the language region of social science) as much as possible. However, because ordinary usage is often ambiguous, encompassing a range of meanings for a given term (as we have seen above for “case”), some concept reconstruction is unavoidable. At the end of this discussion, I hope it will be clear why this particular way of defining terms might be useful, at least for methodological purposes.<sup>11</sup>

*Case* connotes a spatially delimited phenomenon (a unit) observed at a single point in time or over some period of time. It comprises the type of phenomenon that an inference attempts to explain. Thus, in a study that attempts to elucidate certain features of nation-states, cases are comprised of nation-states (across some temporal frame); in a study that attempts to explain the behavior of individuals, cases are comprised of individuals, and so forth. Each case may provide a single observation or multiple (within-case) observations.

For students of political science, the archetypal case is the dominant political unit of our time, the nation-state. However, this is a matter of convention. The study of smaller social and political units (regions, cities, villages, communities, social groups, families) or specific institutions (political parties, interest groups, businesses) is equally common in many social science disciplines.<sup>12</sup> In psychology, medicine, and social work the notion of a case study is usually linked to clinical research, where individuals are the preferred units of analysis.<sup>13</sup> Whatever one’s chosen unit, the methodological issues attached to the case study have nothing to do with the size of the cases. A case may be created out of any phenomenon so long as it has identifiable boundaries and comprises the primary object of an inference.

Note that the spatial boundaries of a case are often more apparent than its temporal boundaries. We know, more or less, where a country begins and ends, while we may have difficulty explaining *when* a country

<sup>11</sup> In the following analysis, I take a “minimal” approach to definition (Gerring 2001: Chapter 4; Gerring and Barresi 2003). Scholars embedded in a particular research setting may choose somewhat different terms and meanings.

<sup>12</sup> For discussion of subnational studies in political science, see Snyder (2001).

<sup>13</sup> Corsini (2004); Davidson and Costello (1969); Hersen and Barlow (1976); Franklin, Allison, and Gorman (1997); Robinson (2001). For discussion of the meaning of the term “case study,” see Benbasat, Goldstein, and Mead (1987: 371); Cunningham (1997); Merriam (1988); and Verschuuren (2003).

begins and ends. Yet some temporal boundaries must be assumed. This is particularly important when cases consist of discrete events – crises, revolutions, legislative acts, and so forth – within a single unit. Occasionally, the temporal boundaries of a case are more obvious than its spatial boundaries. This is true when the phenomena under study are eventful but the unit undergoing the event is amorphous. For example, if one is studying terrorist attacks it may not be clear how the spatial unit of analysis should be understood, but the events themselves may be well bounded.

A *case study* may be understood as the intensive study of a single case where the purpose of that study is – at least in part – to shed light on a larger class of cases (a population). *Case study research* may incorporate several cases, that is, multiple case studies. However, at a certain point it will no longer be possible to investigate those cases intensively. At the point where the emphasis of a study shifts from the individual case to a sample of cases, we shall say that a study is *cross-case*. Evidently, the distinction between case study and cross-case study is a matter of degree. The fewer cases there are, and the more intensively they are studied, the more a work merits the appellation “case study.” Even so, this proves to be a useful distinction, and much follows from it. Indeed, the entire book rests upon it. All empirical work may be classified as either case study (comprising one or a few cases) or cross-case study (comprising many cases).

An additional implication of the term “case study” is that the unit(s) under special focus is not perfectly representative of the population, or is at least questionable. Unit homogeneity across the sample and the population is not assured. If, for example, one is studying a single H<sub>2</sub>O molecule, it may be reasonable to assume that the behavior of that molecule is identical to that of all other H<sub>2</sub>O molecules. Under the circumstances, one would not refer to such an investigation as a “case study,” regardless of how intensive the investigation of that single molecule might be. In social science settings one rarely faces phenomena of such consistency, so this is not an issue of great practical significance. Nonetheless, intrinsic to the concept is an element of doubt about the bias that may be contained in a sample of one or several.

A few additional terms may now be formally defined.

An *observation* is the most basic element of any empirical endeavor. Conventionally, the number of observations in an analysis is referred to with the letter N. (Confusingly, N may also be used to designate the number of cases in a study, a usage that is usually clear from context.) A single observation may be understood as containing several dimensions, each of which may be measured (across disparate observations) as a *variable*.

Where the proposition is causal, these may be subdivided into dependent (Y) and independent (X) variables. The dependent variable refers to the outcome of an investigation. The independent variable refers to the explanatory (causal) factor, that which the outcome is supposedly dependent on.

A case may consist of a single observation (N = 1). This would be true, for example, in a cross-sectional analysis of multiple cases. In a case study, however, the case under study always provides more than one observation. These may be constructed diachronically (by observing the case or some subset of within-case units over time) or synchronically (by observing within-case variation at a single point in time), as discussed below.

This is a clue to the fact that case studies and cross-case studies usually operate at different levels of analysis. The case study is typically focused on within-case variation (if there is a cross-case component, it is probably secondary in importance to the within-case evidence). The cross-case study, as the name suggests, is typically focused on cross-case variation (if there is also within-case variation, it is probably secondary in importance to the cross-case evidence). They have the same object in view – the explanation of a population of cases – but they go about this task differently.

A *sample* consists of whatever cases are subjected to formal analysis; they are the immediate subject of a study or case study. (Confusingly, the term “sample” may also refer to the observations under study. But at present, we treat the sample as consisting of cases.) In a case study, the sample is small, by definition, consisting of the single case or handful of cases that the researcher has under her lens. Usually, however, when one uses the term “sample” one is implying that the number of cases is large. Thus, “sample-based work” will be understood as referring to large-N cross-case methods – the opposite of case study work. To reiterate, the feature distinguishing the case study format from a sample-based (or “cross-case”) research design is the number of cases falling within the sample – one or a few versus many – and the corresponding thoroughness with which each case is studied. Case studies, like large-N samples, seek to represent, in all ways relevant to the proposition at hand, a population of cases. A series of case studies might therefore be referred to as a sample if they are relatively brief and relatively numerous; it is a matter of emphasis and of degree. The more case studies one has, the less intensively each one is studied, and the more confident one is in their representativeness (of some broader population), the more likely one is to describe them as a sample rather than as a series of case studies. For practical reasons – unless, that is, a study is extraordinarily long – the case study research

format is usually limited to a dozen cases or fewer. A single case is not unusual.

Granted, in some circumstances a single study may combine the two elements – an intensive case study and a more superficial analysis conducted on a larger sample. These additional cases are often brought into the analysis in a peripheral way – typically, in an introductory or concluding section of the paper or the book. Often, these peripheral cases are surveyed through a quick reading of the secondary literature or through a statistical analysis. Sometimes, the status of these informal cases is left implicit (they are not theorized as part of the formal research design). This may be warranted in circumstances where the relevant comparison or contrast between the formal case(s) under intensive study and the peripheral cases is obvious. Thus, studies of American exceptionalism, in enumerating features of the American experience, often assume that the United States is different from European countries in relevant respects.<sup>14</sup> In this situation, the additional cases – the UK, France, Germany, and so on – provide the necessary background for whatever arguments are being made about America. They are present, in the sense that they carry an important burden in the analysis, but perhaps they are not formally accounted for in the author's research design. For our purposes, what is significant is that most works combine case study and cross-case study components, whether or not the latter are explicit. Methodologically, these approaches are distinct, even though they may be integrated into a single work. (Indeed, this is a good way of approaching many subjects.)

Continuing with our review of key terms, the sample of cases (large or small) rests within a *population* of cases to which a given proposition refers. The population of an inference is thus equivalent to the breadth or scope of a proposition. (I use the terms *proposition*, *hypothesis*, *inference*, and *argument* interchangeably.) Note that most samples are not exhaustive; hence the use of the term "sample," referring to *sampling* from a larger population. Occasionally, however, the sample equals the population of an inference; all potential cases are studied.

For those familiar with the rectangular form of a dataset, it may be helpful to conceptualize observations as rows, variables as columns, and cases as either groups of observations or individual observations. Several possibilities are illustrated in the tables presented here: two cases (Table 2.1), multiple cross-sectional cases (Table 2.2), and time-series cross-sectional cases (Table 2.3).

<sup>14</sup> Amenta (1991).

TABLE 2.1. Case study dataset with two cases

	X <sub>1</sub>	X <sub>2</sub>	Y
Case 1	Obs 1.1		
	Obs 1.2		
	Obs 1.3		
	Obs 1.4		
	Obs 1.5		
	Obs 1.6		
	Obs 1.7		
	Obs 1.8		
	Obs 1.9		
	Obs 1.10		
	Obs 1.11		
	Obs 1.12		
	Obs 1.13		
	Obs 1.14		
	Obs 1.15		
	Obs 1.16		
	Obs 1.17		
	Obs 1.18		
	Obs 1.19		
	Obs 1.20		
Case 2	Obs 2.1		
	Obs 2.2		
	Obs 2.3		
	Obs 2.4		
	Obs 2.5		
	Obs 2.6		
	Obs 2.7		
	Obs 2.8		
	Obs 2.9		
	Obs 2.10		
Obs 2.11			
Obs 2.12			
Obs 2.13			
Obs 2.14			
Obs 2.15			
Obs 2.16			
Obs 2.17			
Obs 2.18			
Obs 2.19			
Obs 2.20			

Population = 1; Sample = 1; Cases = 2; Observations (N) = 40; Variables = 3.

TABLE 2.2. Cross-case cross-sectional dataset with forty cases

	$X_1$	$X_2$	$Y$
Case 1			
Case 2			
Case 3			
Case 4			
Case 5			
Case 6			
Case 7			
Case 8			
Case 9			
Case 10			
Case 11			
Case 12			
Case 13			
Case 14			
Case 15			
Case 16			
Case 17			
Case 18			
Case 19			
Case 20			
Case 21			
Case 22			
Case 23			
Case 24			
Case 25			
Case 26			
Case 27			
Case 28			
Case 29			
Case 30			
Case 31			
Case 32			
Case 33			
Case 34			
Case 35			
Case 36			
Case 37			
Case 38			
Case 39			
Case 40			

Population

Sample

Population = 1; Sample = 1; Cases = 40; Observations (N) = 40; Variables = 3.

TABLE 2.3. Time-series cross-sectional dataset

	$X_1$	$X_2$	$Y$
Case 1	Obs 1.1	(T <sub>1</sub> )	
	Obs 1.2	(T <sub>2</sub> )	
	Obs 1.3	(T <sub>3</sub> )	
Case 2	Obs 2.1	(T <sub>1</sub> )	
	Obs 2.2	(T <sub>2</sub> )	
	Obs 2.3	(T <sub>3</sub> )	
	Obs 2.4	(T <sub>4</sub> )	
	Obs 2.5	(T <sub>5</sub> )	
Case 3	Obs 3.1	(T <sub>1</sub> )	
	Obs 3.2	(T <sub>2</sub> )	
	Obs 3.3	(T <sub>3</sub> )	
	Obs 3.4	(T <sub>4</sub> )	
Case 4	Obs 4.1	(T <sub>1</sub> )	
	Obs 4.2	(T <sub>2</sub> )	
	Obs 4.3	(T <sub>3</sub> )	
	Obs 4.4	(T <sub>4</sub> )	
	Obs 4.5	(T <sub>5</sub> )	
Case 5	Obs 5.1	(T <sub>1</sub> )	
	Obs 5.2	(T <sub>2</sub> )	
	Obs 5.3	(T <sub>3</sub> )	
	Obs 5.4	(T <sub>4</sub> )	
	Obs 5.5	(T <sub>5</sub> )	
Case 6	Obs 6.1	(T <sub>1</sub> )	
	Obs 6.2	(T <sub>2</sub> )	
	Obs 6.3	(T <sub>3</sub> )	
	Obs 6.4	(T <sub>4</sub> )	
	Obs 6.5	(T <sub>5</sub> )	
Case 7	Obs 7.1	(T <sub>1</sub> )	
	Obs 7.2	(T <sub>2</sub> )	
	Obs 7.3	(T <sub>3</sub> )	
	Obs 7.4	(T <sub>4</sub> )	
	Obs 7.5	(T <sub>5</sub> )	
Case 8	Obs 8.1	(T <sub>1</sub> )	
	Obs 8.2	(T <sub>2</sub> )	
	Obs 8.3	(T <sub>3</sub> )	
	Obs 8.4	(T <sub>4</sub> )	
	Obs 8.5	(T <sub>5</sub> )	

Population

Sample

Population = 1; Sample = 1; Cases = 8; Observations (N) = 40; Time (T) = 1-5; Variables = 3.

It must be appreciated that all these terms are definable only by reference to a particular proposition and a corresponding research design. A country may function as a case, an observation, or a population. It all depends upon what one is arguing. In a typical cross-country time-series regression analysis, cases are countries and observations are country-years.<sup>15</sup> However, shifts in the level of analysis of a proposition necessarily change the referential meaning of all terms in the semantic field. If one moves down one level of analysis, the new population lies within the old population, the new sample within the old sample, and so forth. Population, case, and observation are nested within each other. Since most social science research occurs at several levels of analysis, these terms are generally in flux. Nonetheless, they have distinct meanings within the context of a single proposition and its associated research design.

Consider a survey-based analysis of respondents within a single country, under several scenarios. Under the first scenario, the proposition of interest pertains to individual-level behavior. It is about how individuals behave. As such, cases are defined as individuals, and this is properly classified as a *cross-case* study. Now, let us suppose that the researcher wishes to use this same survey-level data drawn from a single country to elucide an inference pertaining to countries, rather than to individuals. Under this scenario, each poll respondent constitutes a within-case observation. If there is only one country, or a few countries, under investigation – and the inference, as before, pertains to multiple countries – then this study is properly classified as a *case study*. If many countries are under study (with or without individual-level data), then it is properly classified as a *cross-case* study. Again, the key questions are (a) how many cases are studied and (b) how intensively are they studied – with the understanding that a “case” embodies the unit of concern in the central inference.

To complicate matters further, the status of a work may change as it is digested and appropriated by a community of scholars. A *meta-analysis* is a systematic attempt to integrate the results of individual studies into a quantitative analysis, pooling individual cases drawn from each study into a single dataset (with various weightings and restrictions). The ubiquitous *literature review* or *case study survey* aims at the same objective in a less synoptic fashion. Both statistical meta-analyses and narrative literature reviews assimilate a series of studies, treating them as case studies in some larger project – whether or not this was the intention of the original authors.<sup>16</sup>

<sup>15</sup> See, e.g., Przeworski et al. (2000).

<sup>16</sup> Lipsey and Wilson (2001); Lucas (1974).

### A Typology of Covariational Research Designs

In order to better understand what a case study is, one must comprehend what it is not. The distinctiveness of the case study may be clarified by placing it within a broader set of methodological options. Here, I shall classify research designs according to (a) the number of cases that they encompass (one, several, or many), (b) the kind of X/Y variation that they exploit (spatial or temporal), and (c) the location of that variation (cross-case or within-case). This produces a typology with ten possible cells, as depicted in Table 2.4.

Variations on the case study format occupy five of these ten cells, designated by the shaded regions in Table 2.4. Type 2 represents variation in a single case over time (diachronic analysis). Type 3 represents within-case variation at a single point in time (synchronic analysis). Type 4 combines synchronic and diachronic analysis, and is perhaps the most common approach in case study work. Thus, Robert Putnam's classic study of Italy, *Making Democracy Work*, exploits variation across regions and over time in order to test the causal role of social capital.<sup>17</sup>

It is common to combine several cases in a single study. If the cases are comprised of large territorial units, then this combination may be referred to as the “comparative” method (if the variation of interest is primarily synchronic) or the “comparative-historical” method (if the variation of interest is both synchronic and diachronic).<sup>18</sup> It should be pointed out that these terms are used primarily within the subfield of comparative politics. Other terms, such as “most-similar” and “most-different,” may be used as well. Thus, while a case is always singular, a case study work or research design often refers to a study that includes several cases.

The larger point is that the evidentiary basis upon which case studies rely is plural, not singular. Indeed, there are five possible styles of covariational evidence in a case study. Usually, they are intermingled – different sorts of analysis will be employed at different stages of the analysis – so that it is often difficult to categorize a study as falling neatly into a single cell in Table 2.4.

The bottom half of Table 2.4 lays out various cross-case research designs, where the most important element of the empirical analysis involves comparisons across *many* cases (more than a handful). Cross-case

<sup>17</sup> Putnam (1993).

<sup>18</sup> On the comparative method see Collier (1993); Lipphart (1971, 1975); Przeworski and Teune (1970); Richter (1969); and Smelser (1976). On the comparative-historical method see Mahoney and Rueschmeyer (2003). On the history of the comparative method, a term that harkens back to Bryce (1921), see Lasswell (1931).

TABLE 2.4. Research designs: A covariational typology

Cases	Spatial Variation		Temporal Variation	
	None	Within-case	No	Yes
One	1. [Logically impossible]	3. Single-case study (synchronic)	1. [Logically impossible]	2. Single-case study (diachronic)
Several	Cross-case & within-case	4. Single-case study (synchronic-dichronique)	5. Comparative method	6. Comparative-historical
		7. Cross-sectional	8. Time-series cross-sectional	
Many	Cross-case & within-case	9. Hierarchical	10. Hierarchical time-series	

Note: Shaded cells are case study research designs.

analysis without any explicit temporal component (type 7) is usually classified as cross-sectional, even though a temporal component is simulated with independent variables that are assumed to precede the dependent variable. An example was illustrated in Table 2.2. When an explicit temporal component is included, we often refer to the analysis as time-series cross-sectional (TSCS) or pooled time-series (type 8). This format was illustrated in Table 2.3. When one examines across-case and within-case variation in the same research design, one is said to be employing a hierarchical model (type 9). Finally, when all forms of covariation are enlisted in a single research design, the resulting method may be described as a hierarchical time-series design (type 10).<sup>19</sup>

It bears repeating that I have listed the methods most commonly identified with these research designs not with the intention of distinguishing labels but rather with the intention of illustrating various types of causal

<sup>19</sup> It will be noted that, like most case studies, hierarchical models involve a movement across levels of analysis. However, while a case study moves *down* from the primary level of analysis (to within-case cases), a hierarchical model moves *up*. Thus, if classrooms are the primary unit of analysis in a study, one might employ a hierarchical model to control for the effects of larger cases—schools, districts, regions, and so forth. But one would not employ individual students as cases in such an analysis (not, that is, without changing the unit of analysis for the entire study).

evidence. The classification of a research design always depends upon the particular proposition that a researcher intends to prove. Potentially, each of the foregoing cross-case methods might also be employed in the capacity of a case study. (That is, a case study may enlist cross-sectional, time-series cross-sectional, hierarchical, or hierarchical time-series techniques.) It all depends upon the proposition in question (i.e., what sort of phenomena it is about, and hence what sort of phenomena constitutes “cases”) and on the degree of analytic focus devoted to the individual cases.

### The N Question

Traditionally, the case study has been identified with qualitative methods and cross-case analysis with quantitative methods. This is how Franklin Giddings put the matter in his 1924 textbook, in which he contrasted two fundamentally different procedures:

In the one we follow the distribution of a particular trait, quality, habit or other phenomenon as far as we can. In the other we ascertain as completely as we can the number and variety of traits, qualities, habits, or what not, combined in a particular instance. The first of these procedures has long been known as the statistical method. . . . The second procedure has almost as long been known as the case method.<sup>20</sup>

In the intervening decades, this disjunction has become ever more pronounced: a contrast between “stats” and “cases,” “quant” and “qual.” Those who work with numbers are apt to distrust case study methods, while those who work with narratives are likely to be favorably disposed.

I believe that this distinction is not intrinsic, that is, definitional. What distinguishes the case study method from all other methods is its reliance on evidence drawn from a single case and its attempt, at the same time, to illuminate features of a broader set of cases. It follows from this that the number of observations (N) employed by a case study may be either small or large, and consequently may be evaluated in a qualitative or quantitative fashion.<sup>21</sup>

<sup>20</sup> Giddings (1924: 94). See also Meehl (1954); Rice (1928: Chapter 1); and Stouffer (1941: 349).

<sup>21</sup> This section explains and elaborates on a theme first articulated by Lundberg (1941), followed by Campbell (1975/1988) – itself a revision of Campbell’s earlier perspective (Campbell and Stanley 1963). Historical ballast for this view may be garnered from the field of experimental research in psychology, commonly dated to the publication of Gustav Theodor Fechner’s *Elemente der Psychophysik* in 1860. In this work, Hersen and

In order to see why this might be so, let us consider how a case study of a single event – say, the French Revolution – works. Intuitively, such a study provides an N of 1 (France). If one were to broaden the analysis to include a second revolution (e.g., the American Revolution), it would be common to describe the study as comprising two observations. Yet this is a gross distortion of what is really going on. The event known as the French Revolution provides *at least* two observations, for it will be observed over time to see what changed and what remained the same. These patterns of covariation offer essential empirical clues. They also construct multiple observations from an individual case. So N = 2, at the very least (e.g., before and after a revolution), in a case study of type 2 (in Table 2.4).

If, instead, there is no temporal variation – if, for example, the French Revolution is examined at a single point in time – then the investigation is likely to focus on cross-sectional covariational patterns *within* that case, a case study of type 3 (in Table 2.4). If the primary unit of analysis is the nation-state, then within-case cases might be constructed from provinces, localities, groups, or individuals. The possibilities for within-case analysis are, in principle, infinite. In their pathbreaking study of the International Typographers Union, Lipset, Trow, and Coleman note the variety of within-case evidence, which included union locals, union shops (within each local), and individual members of the union.<sup>22</sup> It is not hard to see why within-case N often swamps cross-case N. This is bound to be true wherever individuals comprise within-case observations. A single national survey will produce a much larger sample than any conceivable cross-country analysis. Thus, in many circumstances case studies of type 3 comprise a larger N than cross-sectional analyses or time-series

Barlow (1976: 2–3) report, Fechner developed “measures of sensation through several psychophysical methods. With these methods, Fechner was able to determine sensory thresholds, just noticeable differences (JNDs) in various sense modalities. What is common to these methods is the repeated measurement of a response at different intensities or different locations of a given stimulus in an individual subject . . . It is interesting to note that Fechner was one of the first to apply statistical methods to psychological problems. Fechner noticed that judgments of [JNDs] in the sensory modalities varied somewhat from trial to trial. To quantify this variation, or ‘error’ in judgment, he borrowed the normal law of error, demonstrated that these ‘errors’ were normally distributed around a mean, which then became the ‘true’ sensory threshold. This use of descriptive statistics anticipated the application of these procedures to groups of individuals at the turn of the century when traits of capabilities were also found to be normally distributed around a mean.” Hersen and Barlow note that Fechner, the pioneer, “was concerned with variability *within* the subject.” See also Queen (1928).

<sup>22</sup> Lipset, Trow, and Coleman (1956: 422).

cross-sectional analyses. For example, a recent review of natural resource management studies found that the N of a study varies inversely with its geographic scope. Specifically, case studies focused on single communities tend to have large samples, since they often employ individual-level observations; cross-case studies are more likely to treat communities as comprising observations, and hence have a smaller N.<sup>23</sup> This is a common pattern.

Evidently, if a case study *combines* temporal and within-case variation, as in case studies of type 4, then its potential N increases accordingly. And if cross-case analysis is added to this, as in the comparative method or the comparative-historical method (types 5 and 6 in Table 2.4), then one realizes a further enlargement in potential observations.

These facts hold true regardless of whether the method is experimental or nonexperimental. It is also true of counterfactual reasoning, which typically consists of four observations – the actual (as it happened) before and after observations, and the before and after observations as reconstructed through counterfactual reasoning (i.e., with an imagined intervention). In short, the case study does not preclude a large N. It simply precludes a large *cross-case* N, by definition. Indeed, many renowned case studies are data-rich and include extensive, and occasionally quite sophisticated, quantitative analysis. Frederic Le Play’s work on working-class families incorporated hundreds of case studies.<sup>24</sup> Robert and Helen Lynd’s study of Muncie, Indiana, featured surveys of hundreds of respondents in “Middletown.”<sup>25</sup> *Yankee City*, another pioneering community study, included interviews with 17,000 people.<sup>26</sup>

What, then, of the infamous N = 1 research design that haunts the imaginations of social scientists everywhere?<sup>27</sup> This hypothetical research design occupies the empty cell in Table 2.4. The cell is empty because it represents a research design that is not logically feasible. A single case observed at a single point in time without the addition of within-case observations offers no evidence whatsoever of a causal proposition. In trying to intuit a causal relationship from this snapshot one would be engaging in a truly random operation, since an infinite number of lines might be drawn through that one data point. I do not think there are any

<sup>23</sup> Potete and Ostrom (2005: 11).

<sup>24</sup> Brooke (1970).

<sup>25</sup> Lynd and Lynd (1929/1956).

<sup>26</sup> Warner and Lunt (1941).

<sup>27</sup> Achen and Snidal (1989); Geddes (1990); Goldthorpe (1997); King, Keohane, and Verba (1994); Lieberman (1985: 107–13; 1992, 1994).



examples of this sort of investigation in social science research. Thus, I regard it as a myth rather than a method.<sup>28</sup>

The point becomes even clearer if we consider the case study in relation to a time-series cross-section (TSCS) research design, as illustrated in Table 2.3. Let us imagine that cases are comprised of countries and that temporal units are years; hence, the unit of analysis is the country-year. In Table 2.3, each case has five observations and thus represents a single country observed over five years (T<sub>1-5</sub>). Now, consider the possibility of constructing a case study from just one of these observations – a single country at a single point in time. This seems an unlikely prospect, unless of course there is significant within-case variation during that year. Perhaps this country, during those twelve months, offers a critical juncture in which the variables of theoretical interest undergo a significant change. Whether the temporal era is short or long (and we can imagine much shorter and much longer temporal periods), the significant feature of most case studies is that they look at periods of change, and these periods of change produce (or are regarded as producing) distinct observations – classically “before” (pre-) and “after” (post-) observations. Alternatively, it may be possible to exploit spatial (cross-sectional) evidence in that country at that particular time – for example, with extensive documentary records or a systematic survey. In these circumstances, one can easily imagine a case study being constructed from a single observation in a time-series cross-section research design. But this can be accomplished only by subdividing the original observation into multiple observations. N is no longer equal to 1.

The skeptical reader may regard this conclusion as a semantic quibble, of little import to the real world of research. If so, she might consider the following quite common research scenario. An ethnographic study provides a thick description, in prose, of a particular setting which is intended to uncover certain features of other settings (not studied). The prose stretches for five hundred pages in a draft manuscript and is rather

<sup>28</sup> The one possible exception is the deviant case that disproves a deterministic proposition. However, the utility of the deviant case rests upon a broader population of cases that lies in the background of a case study focused on a single case. Thus, the N of such a study, I would argue, is greater than one – even if no within-case evidence is gathered. The more important point is perhaps the following. No one has ever conducted a case study analysis that consists of only a single observation. If the point of the case study is to demonstrate that a single case of such-and-such a type exists (perhaps with the goal of falsifying a deterministic proposition), then it is likely to take a good deal of work to establish the facts of that case. This work consists of multiple within-case observations. Again, the N is much higher than one.

repetitive; certain patterns are repeated again and again. In an effort to reduce the sheer volume of descriptive material, as well as to attain a more synthetic analysis, the researcher begins to code the results of her labors into standardized categories: she counts. Has she, by committing the act of numeracy, now converted a case study into some other type of study? (If so, what shall we call it?) Note that the object of her study does not vary, even though the prose is now combined with some form of quantitative analysis, which may be simple or sophisticated. The introduction of statistical analysis does not – should not – disqualify a study as a “case study.”

### The Style of Analysis

To be sure, *non*-case study work is by definition quantitative (“statistical”) in nature. This is so because whenever one is attempting to incorporate a large number of cases into a single analysis, it will be necessary to reduce the evidence to a small number of dimensions. One cannot explore 1,000 cases on their own terms (i.e., in detail). (One might simply accumulate case study after case study in a compendious multivolume work. However, in order to reach any meaningful conclusions about this pile of data it will be necessary to reduce the informational overload, which is why God gave us statistics.)

With case study evidence, the situation is evidently more complicated. Case studies may employ a great variety of techniques – both quantitative and qualitative – for the gathering and analysis of evidence. This is one of the intriguing qualities of case-study research and lends that research its characteristic flexibility. Thus, it seems fair to say that there is an elective affinity between the case study format and qualitative, small-N work, even though the latter is not definitionally entailed. Let us explore why this might be so.

Case study research, by definition, is focused on a single, relatively bounded unit. That single unit may, or may not, afford opportunities for large-N within-case analysis. Within-case evidence is sometimes quite extensive, as when individual-level variation bears upon a group-level inference. But not always.

Consider the following classic studies, each of which focuses on the attitudes and characteristics of American citizens. *The American Voter*, a collaborative effort by Angus Campbell, Philip Converse, Warren Miller, and Donald Stokes, examines public opinion on a wide range of topics that are thought to influence electoral behavior through the instrument

of a nationwide survey of the general public.<sup>29</sup> *The People's Choice*, by Paul Lazarsfeld, Bernard Berelson, and Hazel Gauder, is a longitudinal panel study focusing on 600 citizens living in Erie County, Ohio, who were polled at monthly intervals during the 1940 presidential campaign to determine what influences the campaign may have had on their choice of candidates.<sup>30</sup> *Middletown*, by Robert and Helen Lynd, examines life in a midsized city, including such topics as earning a living, making a home, training the young, using leisure, taking part in religious practices, and taking part in community activities (these are the sections into which the book is divided). The Lynds and their accomplices rely on a great variety of evidence, including in-depth interviews, surveys, direct observation, secondary accounts, registers of books checked out of the library, and so forth.<sup>31</sup> *Political Ideology*, by Robert Lane, attempts to uncover the sources of political values in a subsection of the American public, represented by fifteen subjects who are interviewed intensively by the author. These subjects are male, white, married, fathers, between the ages of twenty-five and fifty-four, working-class and white-collar, native-born, of varying religions, and living in an (unnamed) city on the eastern seaboard.<sup>32</sup>

A summary of some of the methodological features of these four studies is contained in Table 2.5. Note that the first two studies (*The American Voter* and *The People's Choice*) are classified as cross-case and the second pair (*Middletown* and *Political Ideology*) as case studies. What is it that drives this distinction? Clearly, it is not the type of subjects under study (all focus primarily on individuals), the number of observations (which range from small-N to large-N), or the breadth of the population (all purport to describe features of the same country). The style of analysis differs in one respect: only in the case studies does qualitative analysis comprise a significant portion of the research. This, in turn, is a product of the number of cases under investigation. Where hundreds of individuals are being studied at once, there is no opportunity to evaluate cases in a qualitative

<sup>29</sup> Campbell et al. (1960).

<sup>30</sup> Lazarsfeld, Berelson, and Gauder (1948). A larger poll, with 2,000 respondents, was taken initially, as a way of establishing a baseline for the chosen panel of 600. In addition, special attention was paid to those whose vote choice changed during the course of the panel. These might be looked upon as a series of case studies nested within the larger panel study. However, because this sort of analysis plays only a secondary role in the overall analysis, it seems fair to characterize this research design as "cross-case."

<sup>31</sup> Lynd and Lynd (1929/1956).

<sup>32</sup> Lane (1962).

TABLE 2.5. Case study and cross-case study research designs compared

	Study	Subjects	Cases	Largest Sample	Analysis	Population
Cross-case study	<i>The American Voter</i> (Campbell et al., 1960)	Citizens of the United States	1,000+ (individuals)	1,000+	Quant.	Americans
	<i>The People's Choice</i> (Lazarsfeld, 1948)	Citizens of Erie County, Ohio	600 (individuals)	2,000	Quant.	Americans
Case study	<i>Middletown</i> (Lynd and Lynd, 1929/1956)	Citizens of Muncie, IN	150+ (cities)	150+	Qual. & Quant.	American cities
	<i>Political Ideology</i> (Lane, 1962)	Working men of Massachusetts	15 (individuals)	15	Qual.	American working-class

All categories (subjects, cases, analysis, population) refer to the primary inferences produced by the study in question.

manner. By contrast, where a single case (as in *Middletown*) or a small number of cases (as in *Political Ideology*) is under study, qualitative analysis is usually de rigueur — though it may be combined with quantitative analysis (as in *Middletown*).

The reader will notice that subtle differences in the research objective of a study can shift it from one category to another. If, for example, Robert and Helen Lynd decided to treat their surveys as representative of *individuals* in the general public (across the United States), rather than as representative of *cities* in the United States, then *Middletown* would take on the methodological features of *The People's Choice*: it would become a cross-case study. Indeed, this is a plausible reading of some portions of that study.

Importantly, the technique of analysis employed in a case study is not simply a function of the sheer number of within-case observations available in that unit. It is, more precisely, a function of the number of *comparable* observations available within that unit. Consider Robert Lane's intensive interviews. Clearly, lots of "data" was recovered from these lengthy discussions. However, the respondents' answers were not coded so as to conform to standardized variables. Hence, they cannot be handled within a dataset format, usually referred to as a "sample" (although we have occasionally employed this term in a broader sense). Of course, Lane could have chosen to recode these interviews to allow

for a quantitative analysis, reducing the diversity of the original information in order to conform to uniform parameters. It is not clear that much would have been gained by doing so. In the event, his study is limited to qualitative forms of analysis.

This issue is treated at length in a later chapter. For the moment, note the fact that case study research often provides a piece of evidence pertaining to A, another piece of evidence pertaining to B, and a third pertaining to C. There may be many observations (in total), and they may all be relevant to a central causal argument, even though they are not directly comparable to one another. These are referred to in Chapter Seven as *noncomparable* observations.

In summary, large-N cross-case research is quantitative, by definition. This much conforms to usual perceptions. However, case study research may be either qualitative or quantitative, or both, depending upon the sort of within-case evidence that is available and relevant to the question at hand. Consequently, the traditional association of case study work with qualitative methods is correctly regarded as a methodological affinity, not a definitional entailment. It is true sometimes, but not all the time.

## What Is a Case Study Good For?

### *Case Study versus Large-N Cross-Case Analysis*

In Chapter Two, I argued that the case study approach to research is most usefully defined as an intensive study of a single unit or a small number of units (the cases), for the purpose of understanding a larger class of similar units (a population of cases). This was put forth as a minimal definition of the topic.<sup>1</sup> In this chapter, I proceed to discuss the *nondefinitional* attributes of the case study – attributes that are often, but not invariably, associated with the case study method. These will be understood as methodological affinities flowing from our minimal definition of the concept.<sup>2</sup>

The case study research design exhibits characteristic strengths and weaknesses relative to its large-N cross-case cousin. These trade-offs derive, first of all, from basic research goals such as (1) whether the study is oriented toward hypothesis generating or hypothesis testing, (2) whether internal or external validity is prioritized, (3) whether insight into causal mechanisms or causal effects is more valuable, and (4) whether the scope of the causal inference is deep or broad. These trade-offs also hinge on the shape of the empirical universe, that is, on (5) whether the population of cases under study is heterogeneous or homogeneous, (6) whether

<sup>1</sup> My intention was to include only those attributes commonly associated with the case study method that are *always* implied by our use of the term, excluding those attributes that are sometimes violated by standard usage. For further discussion of minimal definitions, see Gerring (2001: Chapter 4); Gerring and Barresi (2003); and Sartori (1976).

<sup>2</sup> These additional attributes might also be understood as comprising an ideal-type (“maximal”) definition of the topic (Gerring 2001: Chapter 4; Gerring and Barresi 2003). Recent evaluations of the strengths and weaknesses of case study research can be found in Flyvbjerg (2004); Levy (2002a); and Verschuere (2001).