6

Using Comparative Methods to Study Diversity

Introduction

Comparative researchers examine patterns of similarities and differences across a moderate number of cases. The typical comparative study has anywhere from a handful to 50 or more cases. The number of cases is often limited because one of the concerns of comparative research is to establish familiarity with each case included in a study. Like qualitative researchers, comparative researchers consider how the different parts of each case—those aspects that are relevant to the investigation—fit together; they try to make sense of each case. Thus, knowledge of cases is considered an important goal of comparative research, independent of any other goal.

While there are many types of comparative research (for example, see Mahoney and Rueschemeyer 2003), the distinctiveness of the comparative approach is clearest in studies that focus on diversity (Ragin 2000). Recall that the qualitative approaches examined in Chapter 5 emphasize commonalities, and the primary focus is on similarities across instances (such as when hospital personnel assess the potential social loss of each dying patient). This concern for commonalities dovetails with an interest in clarifying categories and concepts (such as the concept of potential social loss and the situations in which it is assessed). In comparative research on diversity, by contrast, the category of phenomena that the investigator is studying is usually specified at the outset, and the goal of the investigation is to elucidate and explain the

diversity within a particular set of cases. (This type of comparative research, which is the major focus of this chapter, is examined in detail in Ragin 2000.)

Consider the following example of comparative research on diversity. From the mid-1970s through the 1980s, many less developed countries experienced mass protest in response to "austerity programs" demanded by the International Monetary Fund (IMF). (These mass protests were similar to those that occurred in Greece in May of 2010 after that government reached an agreement with the IMF that included increasing taxes, cutting public spending, and freezing government wages and pensions for three years.) These countries had accumulated large public debts that they could not repay. In exchange for better terms (e.g., lower interest rates and longer repayment periods), the governments of these countries agreed to implement a variety economic policies designed to facilitate debt repayment. For instance, in some countries the IMF demanded that the government stop subsidizing the prices of basic commodities such as fuel and food. These austerity measures saved governments money and made debt repayment more feasible; they also provoked widespread protest among citizens faced with new challenges to their economic and social well-being (Walton and Ragin 1990).

A comparative researcher interested in these countries might contrast the different *forms* of protest that occurred in response to these austerity programs. In some countries, there were riots; in others, there were labor strikes led by unions; in others, there were mass demonstrations involving many different groups; in still others, opposition political parties organized protests; and so on. Why did different kinds of protest erupt in different countries? What causal conditions explain these different responses to austerity programs? And why did some countries with severe austerity programs experience very mild mass protest?

To explain this diversity, a comparative researcher would first group countries according to their different responses to austerity, placing all those with riots in one category, those with demonstrations in another, and so on. Next, the investigator would look for patterns of similarities and differences. What are the similarities among the countries with riots that distinguish them from the other countries? Perhaps the countries with riots also had repressive governments, widespread poverty, and serious crowding in major urban areas, and none of the non-riot countries had this specific combination of conditions. How did the countries with mass demonstrations differ from all the other countries with austerity programs? This systematic search for differences would continue until the researcher could account for the diverse responses to austerity found in these countries.

Thus, in research that emphasizes diversity, the focus is on the similarities within a category of cases with the same outcome (e.g., countries with riots) that (1) distinguish that category from other categories (countries with other forms of austerity protest) and (2) explain the outcome manifested by that category. In other words, the study of diversity is the study of patterns of similarities and differences within a given set of cases (in this example, countries with austerity protests).

Contrasts With Other Research Strategies

As already noted, the main difference between comparative research on diversity and qualitative research on commonalities is that their basic orientation toward cases differs. When qualitative researchers study commonalities, they usually view multiple cases as many instances of the same thing. A qualitative researcher who interviews many taxicab drivers, for example, uses these many instances to deepen the portrayal of this case—the taxicab driver.

Comparative researchers who study diversity, by contrast, tend to look for differences among their cases. Comparative researchers examine patterns of similarity and difference across cases and try to come to terms with their diversity. A comparative researcher might study the settlement of Somali refugees in the United States since the onset of the 1991 Somali Civil War, contrasting the ways they were received in a variety of different communities. It might be possible to distinguish four or five basic types of reception—from hostile to indifferent to open to paternalistic and so on—and then to pinpoint the factors (such as the size and wealth of a community) that determined these different receptions. Another comparative researcher might study bars in a community and contrast the different strategies they use to encourage and discourage drinking (for those patrons who have been over-served). Bars that cater to different customers (for example, bikers versus college students) surely use different techniques. In each of these examples, the research focuses on the diversity that exists within a specific set of cases.

Quantitative researchers (the focus of Chapter 7) also examine differences among cases, but with a different emphasis. In quantitative research, the goal is to explain the covariation of one variable with another, usually across many, many cases. A quantitative researcher, for example, might explain different levels of income across thousands of individuals included in a survey by pointing to the covariation between income levels and educational levels—people with more education tend to have more income. In quantitative research, the

focus is on differences in levels and how different variables like income and education covary across cases. In comparative research, by contrast, the focus is on diversity—*patterns* of similarities and differences. In other words, the focus is on the various configurations of a set of attributes (such as the basic types of reception of Somali refugees in the United States).

Furthermore, a quantitative researcher typically has only broad familiarity with the cases included in a study. As the number of cases exceeds 50 or so, it becomes increasingly difficult to establish familiarity with each case. Imagine a survey researcher trying to become familiar with the lives of the thousands of people included in a telephone survey or a political scientist trying to keep up with major elections in all democratic countries. Neither task is feasible. There are practical limits to how many cases a researcher can study closely.

We emphasize how these three research strategies typically have different approaches toward case selection, use different techniques to collect and analyze evidence, and have different underlying research goals. Yet it is important to reiterate that there is a unity among these strategies, as they all involve the interplay of ideas and (large amounts of) evidence to produce representations. The distinctions we draw between these strategies are not absolute. Consider an example that blurs these distinctions. The 2007 book *Bargaining for Brooklyn: Community Organizations in the Entrepreneurial City,* by Nicole Marwell, is a qualitative study of eight nonprofit, community-based organizations in Brooklyn, New York. By developing in-depth knowledge of her cases, Marwell is able to uncover a range of organizational strategies used by community-based organizations in their attempts to help local residents move out of poverty. Her focus is on the patterns of similarities and differences across these cases, where diversity in organizational strategies emerges even though all the organizations operate in the same neighborhood. (For an example of quantitative researchers attending to diversity, see Grant, Morales, and Sallaz 2009.) As mentioned previously, researchers will use whatever procedures they believe will best illuminate their object of study. Many researchers use multiple strategies to study the same phenomenon. They hope that the weaknesses of one approach will be counterbalanced by the strengths of another (see, for example, Fernandez-Mateo 2007).

The Goals of Comparative Research

The emphases of comparative research on diversity (especially the different patterns that may exist within a specific set of cases) and on familiarity with each case make this approach especially well suited for the goals

of exploring diversity, interpreting cultural or historical significance, and advancing theory.

Exploring Diversity

The comparative approach is better suited for addressing patterns of diversity than either of the other two strategies. Diversity is most often understood in terms of types of cases. The typical goal of a comparative study is to unravel the different causal conditions connected to different outcomes—causal patterns that separate cases into different subgroups. This explicit focus on diversity distinguishes the comparative approach from the qualitative approach. Recall that in qualitative research, the goal is often to clarify categories with respect to the concepts they exemplify by examining similarities across the instances of a category (such as taxicab drivers).

One common outcome of comparative research is the finding that cases that may have been defined as "the same" at the outset are differentiated into two or more categories at the conclusion of the study. For example, a researcher studying major U.S. cities that have elected African American mayors might conclude at the end of the study that there are two major types—those cities where interracial alliances resulted in the successful election of African American mayors and those cities where black voters, who happened also to constitute a majority of voters, made the difference in the election of African American mayors. The political dynamics and significance of the elections could differ considerably across the two types.

A researcher studying governments that terrorize citizens who oppose them might find that there are several main types, depending on the international standing of the government in question. For instance, when a government is supported by the United States and other major powers, its terrorizing may be overlooked. When a government lacks this support, its terrorizing may be considered repugnant. Governments in the second category would have to contend with the possibility that their actions might provoke international sanctions or intervention and therefore would likely practice more covert forms of terror.

While comparative researchers often discern *types* in the course of their examinations of patterns of diversity, they may also begin their research with a tentative delineation of types. A common strategy is to categorize cases according to their different outcomes. The goal of such a strategy is to unravel the causal conditions that generate various outcomes. If different causes can be matched to the different outcomes, then the research confirms the investigator's understanding of the factors that distinguish these cases. If not, then the frame for the research needs to be reformulated.

For example, a researcher might examine the causes of different types of government repression. Some repressive governments, for example, may simply harass their opponents—incarcerating them for short periods, subjecting them to frequent questioning, opening their mail, and so on. Other repressive governments may torture and kill their opponents (as Iranian President Mahmoud Ahmadinejad has done since his election in 2005 [McManus 2009]). Still other governments may focus their repressive energies not on opponents, but on purging the less committed from their own ranks—members of the ruling political party or clique (as did Saddam Hussein by executing hundreds of members of the Ba'ath Party in Iraq [Makiya 1998]. Still others may attack random members of society in order to maintain a general state of terror and obedience (as Joseph Stalin did in the Soviet Union during the 1930s [Gregory 2009]). It is important to understand different types of repression and the various conditions that explain the emergence of each type, particularly for researchers whose fundamental goal is to generate knowledge with the potential to transform society (see Chapter 2).

The goal of exploring diversity is important because people, including social researchers, sometimes have trouble seeing the trees for the forest. That is, they tend to assume uniformity or generality when, in fact, there is a great deal of diversity. Here is a simple example: Generally, governments that are less democratic tend to be more repressive. However, there are many instances of repression by democratically elected governments and many instances of politically tolerant and lenient governments that are not democratic. To understand government repression fully, it is necessary to go beyond the simple identification of political repression with an absence of democracy and examine the different forms of government repression that exist in all countries.

Interpreting Cultural or Historical Significance

Comparative researchers focus explicitly on patterns of similarities and differences across a range of cases. Relevant cases, in turn, are almost always drawn from a specific and known set. Recall that in qualitative research (Chapter 5), much energy is often devoted to coming to terms with the case: What is this case a case of? What concepts are exemplified in this case? Into which larger social scientific categories, if any, does it fit? In comparative research, by contrast, the researcher usually starts with a good sense of the larger category that embraces the cases included in the study because this category is usually specified beforehand (such as "countries with austerity protests").

For example, a researcher might focus on "military coups in Latin America since 1975" or "major cities in the United States that have elected African American mayors." In these examples, the relevant set of cases is defined in advance, and there is a finite, usually moderate number of such cases. Typically, the category that establishes the boundary of the set is historically and geographically delimited. In each of the examples just mentioned, time and place boundaries are either plainly stated ("Latin America since 1975") or implied ("recent U.S.").

This focus on circumscribed categories makes the comparative strategy well suited for the goal of interpreting historically or culturally significant phenomena, especially when there are a moderate number of cases, as in the examples just mentioned. The category "major cities in the United States that have elected African American mayors," for example, is historically significant in part because it is a relatively new and major phenomenon. Prior to the expansion of civil rights in the 1960s, there were no African American mayors in major U.S. cities. It is culturally significant because of the relevance of race and race relations to American society. Likewise, the category "military coups in Latin America since 1975" is significant to those concerned with the progress of democracy and human rights in this region.

Because the comparative approach focuses on differences between cases and the differentiation of types, it facilitates historical interpretation. Consider the category "revolution." Some revolutions simply change those who are in power or alter other political arrangements without implementing any major changes in society. The revolutionaries who overthrew Ferdinand Marcos in the Philippines, for example, did not attempt any fundamental changes in Philippine society. Other revolutions, by contrast, bring with them regimes that seek to alter society fundamentally: Kings are beheaded, property is confiscated, basic social patterns and relations are changed forever. Revolutionary social changes of this nature were attempted after the French Revolution of 1789, the Russian Revolution of 1917, and the Chinese Revolution of 1949.

Revolutions that attempt fundamental social change are treated as a distinct type by social scientists. These massive upheavals of society are called "*social* revolutions" to distinguish them from revolutions that simply change leaders or other political arrangements (Skocpol 1979). By differentiating social revolutions from all other kinds, researchers provide important tools for understanding and interpreting these massive social transformations. When a major upheaval occurs, researchers can assess whether or not it qualifies as a social revolution. If so, it can be compared with other social

revolutions. If not, then some other category may be used (for example, coup d'etat) to interpret the event and to specify comparable cases. Generally, when a set of comparable cases can be specified, these cases aid the interpretation and understanding of the new case.

More generally, when social scientists categorize an event, they establish a primary analytic frame for its interpretation. Thus, the interpretation of historically or culturally significant events is often a struggle over the proper classification of events into broad categories—a key concern of the comparative approach.

Advancing Theory

Several basic features of the comparative approach make it a good strategy for advancing theory. These features include its use of flexible frames, its explicit focus on the causes of diversity, and its emphasis on the systematic analysis of similarities and differences in the effort to specify how diversity is patterned.

In comparative research, investigators usually initiate research with a specific analytic frame, but these initial frames are open to revision. The researcher interested in military coups in Latin America since 1975, for example, already has a frame for the research—the frame of military coups. Recall that the frames of qualitative research are fluid, and researchers may not finish developing their frames until after all the work of collecting and studying the evidence is complete. In comparative research, by contrast, frames are established at the outset of a research project, but they remain flexible. Comparative researchers expect their frames to be revised, and in fact conduct research in order to sharpen the ideas expressed in a frame.

A researcher interested in welfare states in advanced countries, for example, might start out with a frame that specifies two basic types of welfare states but then conclude with a specification of three or four types (Esping-Andersen 1999). Or, the researcher might conclude that there is only one main type and that all deviations are best understood as underdeveloped or incomplete expressions of the main type (J. Stephens 1979). By altering initial frames in response to evidence, comparative researchers refine and elaborate existing ideas and theoretical perspectives.

When conducting their research, comparative researchers are more explicitly concerned with causation and causal complexity than are most qualitative researchers. For example, when comparative researchers differentiate types (such as types of government repression), they also try to specify the combinations of causal conditions conducive to each type: What causes some

regimes to concentrate their repressive efforts on regime opponents? What causes others to focus their efforts on purging troublesome members of the ruling party? And what causes still other regimes to cultivate a general state of terror in the population at large? This emphasis on causation is central to theory because most theories in the social sciences are concerned with explaining how and why—that is, with specifying the causes of social phenomena.

To assess causation, comparative researchers study how diversity is patterned. They compare cases with each other and highlight the contrasting effects of different causes. Comparative researchers view each case as a combination of characteristics (for example, conditions relevant to government repression) and examine similarities and differences in combinations of characteristics across cases in their effort to find patterns.

The Process of Comparative Research

The comparative study of diversity is neither as fluid as qualitative research nor as fixed as quantitative research. Comparative researchers typically start with a carefully specified category of phenomena that is intrinsically interesting in some way (for example, countries with austerity protests). They use analytic frames to help them make sense of their categories, and they revise their frames based on their examination of evidence.

In the course of their research, they focus on patterns of similarities and differences among cases and assess patterns of diversity. This assessment of diversity provides the foundation for improving or revising the analytic frame chosen at the outset of the study. Like qualitative research, the comparative approach stimulates a rich dialogue between ideas and evidence. Researchers generate images from their data and adjust their frames as they construct representations of their research subjects.

Selecting Cases

Comparative researchers usually initiate their research with a specific set of cases in mind. Most often, this set has clear spatial and temporal boundaries and embraces cases that are thought to be comparable with each other, as in the examples already described. The degree to which the cases that are selected actually belong to the same category (and therefore are comparable) is assessed in the course of the research. While conducting the investigation, the researcher may decide that some cases don't belong in the same category as the others and can't be compared. They also may reformulate the category as the research proceeds. Usually, however, such adjustments are modest.

Typically, the cases that comparative researchers select for study are specific to their interests and to those of their intended audience—for example, countries with mass protest against IMF-mandated austerity programs. This category of countries has clear spatial and temporal boundaries and embraces a set of comparable cases. It is also an intrinsically interesting set of cases. In short, it is just the kind of delimited empirical category that is well suited for comparative investigation.

The comparative approach can be applied to many different kinds of cases, not just countries. It is important, however, for the cases selected to be comparable and to share membership in a meaningful, empirically defined category. For example, the comparative approach can be applied to the fraternities on a college campus, to refugee groups living in a major urban area, to different religious congregations in a medium-sized town, to the truck stops along Interstate 55, or to the elections in the congressional districts of a large state. The set of cases must be coherent. Usually, the cases must also offer some potential for advancing social scientific thinking.

Using Analytic Frames

When researchers choose their cases, they also usually select their analytic frames. Essentially, a frame is chosen when the researcher specifies what about the cases is of interest. For example, the researcher wanting to study countries with austerity protests may be interested in the different forms that the protest took. This frame, which would be developed from the existing social science literature on mass protest, would specify how people respond to different conditions in different ways when they engage in political protest. In short, it would detail the different kinds of factors the researcher should examine in a comparative study of protest.

To continue this example, in some countries, opposition groups may have many resources; in others, they may have few. Groups with more resources are more likely to engage in organized activities such as strikes and in other activities that are relatively costly to participants. People on strike, for example, must give up their wages. Thus, this frame, which would be developed from the existing literature on social movements and collective action, would direct the researcher to focus on resource mobilization, among other things. Analytic frames help researchers see aspects of cases that they might otherwise overlook, and direct their attention away from other aspects.

Sometimes researchers are interested in many facets of their cases and don't select a frame until they are well along in their research. It might take awhile, for example, to determine what a comparison of countries with

austerity protests might best offer in the way of general social scientific knowledge. Comparative researchers also may develop new frames from their evidence—for example, a new frame for the study of race and politics based on a study of cities where coalitions of white and black voters have elected African American mayors. This practice is less common in comparative research than in qualitative research, however, because comparative researchers start with a fairly good sense of their cases and the empirical category that embraces them (such as "countries that experienced mass protest in response to IMF-mandated austerity").

Analyzing Patterns of Diversity

In comparative research, the examination of diversity—patterns of similarities and differences—goes hand-in-hand with the study of causes. Generally, researchers expect different causal conditions to be linked to divergent outcomes in interpretable ways. Thus, the goal of the researcher's examination of patterns of similarities and differences is to identify causal links—how different **configurations** of causes produce different outcomes across the range of cases included in a study. The specification of different patterns of causation is the primary basis for the differentiation of types.

In a study of how sororities generate a feeling of group solidarity, for example, different ways of generating this feeling should affect the *nature* of the solidarity that is generated. The researcher might find that some sororities generate solidarity around special events and rituals, while others generate it through routine activities that bring members of the sorority together on a daily basis. These different ways of generating this feeling (that is, these different causes of solidarity) should have consequences for the nature of the solidarity observed. For example, solidarity in sororities of the first type may be more visible but also less durable, while in the second type, it may be more subtle but more enduring.

If causes and outcomes cannot be linked in interpretable ways, then researchers must reexamine their specification of causes and outcomes and their differentiation of types. In many ways, this process of differentiating types and specifying causal links specific to each type resembles the "double fitting" of categories and images that constitutes the core of qualitative methods (see Chapter 5). There is a dialogue between ideas and evidence that culminates in a meaningful representation of the research subject. The main difference is that in qualitative research, the emphasis is on clarifying a category and enriching its representation, whereas in comparative research, the emphasis is on using contrasts among cases to further the researcher's understanding of their diversity.

Using Comparative Methods

Comparative methods are used to study configurations. A **configuration** is a specific combination of attributes that is common to a number of cases. For example, if all the countries that experienced mass demonstrations in response to IMF-mandated austerity were similar in having low levels of economic development, high levels of urbanization, undemocratic governments, and poorly organized opposition groups, this would constitute a specific configuration of conditions associated with mass demonstrations as a response to IMF-mandated austerity. The examination of patterns of diversity essentially involves a search for combinations of conditions that distinguish categories of cases. Thus, researchers look for uniformity *within* categories and contrasts *between* categories in combinations of conditions.

Data procedures appropriate for the study of configurations, initially formalized by Drass and Ragin (1989), constitute the core of the comparative approach to diversity. Comparative methods are used to examine complex patterns of similarities and differences across a range of cases. Like quantitative methods (see Chapter 7), comparative methods are used to examine causes and effects, but the emphasis in comparative research is on the analysis of configurations of causal conditions.

Before examining data procedures specific to comparative methods, let's first consider an example that shows the main ideas behind the techniques.

An Overview of Comparative Methods

A hypothetical example, based on a study of the *repression* of austerity protests conducted by Walton and Ragin (1990), is used to illustrate general features of comparative methods. Table 6.1 presents hypothetical data on 16 countries that experienced austerity protests in the early 1980s. Eight of these countries had governments that became violently repressive in response to austerity protests; the governments of the other eight did not, even though they also experienced such protests.

The table shows differences and similarities among these 16 countries with respect to conditions believed to be relevant to repression, derived from an analytic frame for government repression. The conditions include the following:

1. Whether the country was politically aligned with the Soviet Union or with the United States and Western Europe in the 1980s

2. Whether or not the country had undergone substantial industrialization prior to 1980

3. Whether or not the country had a democratic government prior to the emergence of austerity protests

4. Whether or not the country had a strong military establishment prior to the emergence of austerity protests

The goal of comparative analysis is to determine the combinations of causal conditions that differentiate sets of cases. In this analysis, the goal is to find combinations of causal conditions that distinguish the eight countries with governments that became repressive from the other eight countries. Careful examination of the similarities among the countries with violently repressive governments shows that they do not share any single causal condition or any single combination of conditions. However, there are two combinations of conditions that are present in the set of countries that had repressive governments that are both absent from the set that did not. The 16 cases are sorted in the table to highlight these two combinations.

The first four cases share an absence of democratic government prior to the emergence of austerity protests combined with a strong military establishment. None of the cases in the lower half of the table (the eight countries lacking violent repression) has this combination. The second four countries with violent repression share two different conditions: a presence of democratic government prior to austerity protests combined with an absence of significant industrialization prior to the protests. Again, none of the eight countries lacking violent repression has this combination of conditions.

The results of the examination of similarities and differences thus lead to the conclusion that there are two different combinations of conditions (or causal configurations) that explain the emergence of violent repression in these cases. The first configuration (nondemocratic rule combined with a strong military) suggests a situation where the military establishment has gained the upper hand in part because of the absence of checks (democratic government) on its power. The second configuration (absence of significant industrialization combined with presence of democratic government prior to the emergence of violent repression) suggests a situation where a breakdown of democratic rule occurred in countries that lacked many of the social structures associated with industrialization (e.g., urbanization, literacy, and so on). These social structures are believed to facilitate stable democratic rule. Further research might show important differences between these two sets of cases with respect to the kind of repression that was inflicted on the protesters.

Table 6.1   Simple Example of Comparative Methods*

Case	Aligned With USSR	Industrialized	Democratic Government	Strong Military	Violent Repression[†]
1	0	0	0	1	1
2	0	1	0	1	1
3	1	0	0	1	1
4	1	1	0	1	1
5	0	0	1	0	1
6	0	0	1	1	1
7	1	0	1	0	1
8	1	0	1	1	1
9	0	0	0	0	0
10	0	1	0	0	0
11	0	1	1	0	0
12	0	1	1	1	0
13	1	0	0	0	0
14	1	1	0	0	0
15	1	1	1	0	0
16	1	1	1	1	0

*In the columns with causal or outcome conditions, the number 1 indicates the presence of a condition, or "yes"; 0 indicates its absence, or "no."
[†]The two combinations of conditions linked to violent repression are (1) absence of democratic government combined with strong military and (2) presence of democratic government combined with an absence of industrialization.

The cases are arranged in Table 6.1 so that the main patterns of similarity among the countries with violent repression are easy to detect, and the comparison of these cases with countries lacking violent repression is simplified. Specific procedures for assessing patterns of similarity and difference are detailed in the next section. Before examining these procedures, consider several general features of the comparative analysis just presented:

- Comparative analysis proceeds by comparing configurations of causes—*rows* of the table—and not by comparing the presence or absence of each causal condition (that is, each of the first four columns) with presence or absence of the outcome (the last column—repression).

- Comparative analysis allows for the possibility that there may be several combinations of conditions that generate the same general outcome (government repression in the example).

- Comparative analysis can address complex and seemingly contradictory patterns of causation. One causal condition (democratic government prior to the emergence of violent repression) is important in both its present and absent condition—it appears in both causal configurations, but contributes in opposite ways.

- Comparative analysis can eliminate irrelevant causes. One causal condition (whether the country was aligned with the Soviet Union) was eliminated as an important causal condition. Even though it was considered a possible factor at the outset, examination of similarities and differences among repressive and nonrepressive cases shows that this cause is not an essential part of either of the key causal combinations.

The findings in Table 6.1 are easy to see. Usually, however, the patterns are not so simple, and researchers must use more systematic comparative methods to help them analyze similarities and differences. These techniques, explained in the next sections, make it possible for researchers to find patterns that they would probably miss if they tried to unravel differences simply by "eyeballing" their cases.

Specifying Causes and Outcomes

In the comparative approach, each case is understood as a combination of causal conditions linked to a particular outcome. Thus, the selection of the outcome to be studied and the specification of causal conditions relevant to that outcome are crucially important parts of a comparative investigation.

Generally, in order to specify causes, the investigator must be familiar with the research literature on the outcome (for example, "government repression" in the study just described) and with the cases included in the study. In this early phase of the research, the investigator explores connections between social scientific thinking (for example, about government repression) and the

evidence. These early explorations lead to a clarification of the nature of the outcome to be studied and a specification of the relevant causes.

The comparative methods described in this chapter use what social scientists call **presence–absence dichotomies.** This means that causal conditions and outcomes are either present or absent in each case and can be coded "yes" or "no," as in Table 6.1. Thus, instead of using a precise measure of industrialization (for example, the percentage of the work force employed in manufacturing) in the data analysis, an assessment is made of whether or not substantial industrialization has occurred. The use of presence–absence dichotomies simplifies the representation of cases as configurations of causal conditions. It is important to note, however, that comparative analysis is not limited to presence–absence dichotomies. In fact, any type of measure can be used in a comparative analysis. As Ragin shows in *Redesigning Social Inquiry* (2008), researchers can examine configurations of causal conditions that vary by level or degree. The key is to use measures that scale the degree of membership in sets, using scores that range from 0 to 1, instead of simply absent (0) or present (1)— the schema used in this chapter to illustrate the logic of comparative research. The use of measures that vary by level or degree is discussed as well in Chapter 7; however, Chapter 7 focuses on the correlations between measures, not on different configurations of conditions.

When using causal conditions expressed in presence–absence dichotomies, the number of conditions determines the number of causal combinations that are possible. For example, the specification of four causal conditions (as in the middle four columns of Table 6.1) provides for 16 (that is, 2^4) logically possible combinations of causal conditions (all 16 appear in Table 6.1). Specification of five causal conditions provides for 32 (2^5) combinations, six causal conditions provides for 64 (2^6) combinations, and so on. Causal conditions are not examined separately, as in studies focusing on covariation, but in combinations.

Once causal conditions have been selected, cases conforming to each combination of causal conditions are examined to see if they agree in terms of the outcome. In Table 6.1, there is only one case for each combination of causal conditions, so there is no possibility of disagreement. But what if there were two cases in the first row (that is, two countries that combined absence of alignment with the Soviet Union, absence of substantial industrialization before 1980, absence of democratic government, and presence of a strong military), but in one country protesters suffered violent repression and in the other they did not? The researcher would have to determine what additional factor (present in one country but absent in the other)

caused repression. This new causal condition would then be added to the table for all cases.

If there are many causal combinations with cases that disagree on the outcome, then the investigator should take this as a sign that the specification of causal conditions is incorrect or incomplete. The close examination of cases that have the same presence–absence values on all the causal conditions yet have different outcomes is used as a basis for selecting additional causal variables. Investigators move back and forth between specification of causal conditions (using social science theory and their general substantive knowledge as guides) and examination of evidence to resolve these differences.

Constructing the Truth Table

Once a satisfactory set of causal conditions for a particular outcome has been identified, evidence on cases can be represented in **truth tables.** A truth table summarizes a **data matrix** (e.g., Table 6.1) by sorting cases according to their combinations of values on dichotomous (yes/no) causal variables. The use of truth tables facilitates the analysis of patterns of similarities and differences.

The first step in constructing a truth table is simply to list the evidence on the cases in the form of a data table. Consider, for example, the data presented in Table 6.2. This table shows hypothetical evidence on 30 suburban school districts surrounding a major metropolitan area. The outcome of interest here is whether or not the elementary schools in each district track students according to ability. When students are *tracked*, they are grouped together into relatively homogeneous classes. Students who learn things quickly are assigned to one class, while students who learn things at an average speed are assigned to another, and so on.

Having students of uniform ability together in the same room is thought to simplify teaching, making it more efficient. After all, it clearly would be a mistake to put first graders and eighth graders in the same classroom. Why not apply this same principle to students within grade levels? The usual objection is that students who are assigned to the "slow" group become branded low achievers and are rarely given the opportunity to prove otherwise. Plus, being surrounded by "faster" students can motivate a "slow" student to learn faster. Assigning students to the slow group may seal their academic fate.

The researcher in this example wanted to understand why some school districts track elementary school students and others don't. The middle columns

Table 6.2   Hypothetical Data on Tracking in School Districts[*]

School District	Racial Diversity	Class Diversity	Competitive Elections	Unionized Teachers	Ability Tracking
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	1	1
5	0	0	0	1	1
6	0	0	1	0	0
7	0	0	1	1	1
8	0	1	0	0	0
9	0	1	0	0	0
10	0	1	0	0	0
11	0	1	0	0	0
12	0	1	0	1	1
13	0	1	1	0	0
14	0	1	1	1	1
15	1	0	0	0	1
16	1	0	0	0	1
17	1	0	0	1	1
18	1	0	0	1	1
19	1	0	0	1	1
20	1	0	0	1	1
21	1	0	1	0	0
22	1	0	1	0	0
23	1	0	1	0	0
24	1	0	1	1	0
25	1	1	0	0	1
26	1	1	0	1	1
27	1	1	0	1	1
28	1	1	1	0	0
29	1	1	1	1	0
30	1	1	1	1	0

[*]In the columns with causal or outcome conditions, the number 1 indicates the presence of a condition, or "yes"; 0 indicates its absence, or "no."

of the table list the causal conditions that the researcher, based on an examination of the relevant research literatures, thought might be important:

1. Whether the school district is racially diverse or predominantly white

2. Whether or not the school district has a broad representation of income groups (poor, working class, middle class, and upper middle class)

3. Whether or not the school board elections in the district are open and competitive, with good voter turnout

4. Whether or not the teachers in the district are unionized

The first two factors (racial and class diversity) show the social composition of school districts. These factors are important because where there is more diversity, members of dominant groups (for example, whites in racially diverse districts) generally believe that tracking will benefit their children most. The competitiveness of school board elections is important because the majority of voters usually disapprove of tracking in elementary schools. They believe this practice benefits only a small number of students. In districts where school board elections are routine matters that attract little voter interest, however, the small number of families that benefit from tracking might have more influence. Unionization of teachers is included because the researcher believes that teacher unions prefer tracking because it simplifies teaching.

The school districts are sorted in Table 6.2 according to the four causal conditions so that districts that are identical on these factors are next to each other. Inspection of the data shows that there are no districts that have the same combination of scores on the causal conditions but different outcomes. Districts 8–11, for example, all show the same pattern on the four causal conditions; they also are identical on the outcome—none of these districts tracks students according to ability. If the cases were not consistent on the outcome, it would be necessary to examine them closely to determine which other causal factors should be added to the table.

Listing the data on the cases, as shown in Table 6.2, is a necessary preliminary to the construction of the truth table. The idea behind a truth table is simple: The focus is on causal combinations. Each logical combination of values on the causal conditions is represented as one row of the truth table. Thus, truth tables have as many rows as there are logically possible combinations of values on the causal conditions. If there are four dichotomous causal conditions, as in Table 6.2, the truth table will contain $2^4 = 16$ rows. Each row of the truth table is assigned an outcome score (1 or 0, for presence or absence of the outcome) based on the cases in that row. The first three cases in Table 6.2, for example, have the same combination of scores on the causal conditions (absent on each of the four conditions) and the same outcome (absence of tracking). They are

combined to form the first row of the truth table presented in Table 6.3. The number of districts that make up each row of the truth table is also reported in Table 6.3, so that the translation of Table 6.2 to Table 6.3 is clear.

Simplifying the Truth Table

The truth table (Table 6.3) summarizes the causal configurations that exist in a data table (Table 6.2). Listing configurations is not the same as identifying patterns, however. Usually, comparative researchers want to examine configurations to see if they can be simplified. When investigators simplify configurations, they identify patterns.

Here is a quick example of simplification: Look at rows 13 and 14 of the truth table reported in Table 6.3. Row 13 reports that school districts that combine the following four characteristics track students: (1) racial diversity, (2) class diversity, (3) an absence of competitive school board elections, and (4) an absence of teachers' unions. Row 14 reports that school districts that differed on only one of these four conditions—districts in this row had teachers' unions—also tracked students. The comparison of these two rows shows that when the first two causal conditions are present (race and class diversity) and the third is absent (competitive school board elections), it does not matter whether or not teachers are unionized; tracking by ability still takes place.

An easy way to represent this simplification is to use uppercase letters to indicate presence of a condition and lowercase letters to indicate its absence. In this example, the word RACE indicates the presence of racial diversity; the lowercase word race is used to indicate its absence. Similarly, the word CLASS is used to indicate the presence of class diversity; the lowercase class is used to indicate its absence. ELECTIONS is used to indicate the presence of open, competitive school board elections; elections is used to indicate the absence of this condition. UNIONS is used to indicate the presence of teachers' unions; unions is used to indicate the absence of this condition. Finally, TRACKING is used to indicate the presence of tracking, and tracking is used to indicate its absence.

Thus, row 13 can be represented as

TRACKING = RACE·CLASS·elections·unions

Row 14 can be shown as

TRACKING = RACE·CLASS·elections·UNIONS

where the multiplication symbol (·) is used to indicate the combination of conditions. These two rows can be simplified through combination

Table 6.3	Truth Table for Data on Tracking in School Districts*

Row	Racial Diversity	Class Diversity	Competitive Elections	Unionized Teachers	Ability Tracking	Number of Districts†
1	0	0	0	0	0	3
2	0	0	0	1	1	2
3	0	0	1	0	0	1
4	0	0	1	1	1	1
5	0	1	0	0	0	4
6	0	1	0	1	1	1
7	0	1	1	0	0	1
8	0	1	1	1	1	1
9	1	0	0	0	1	2
10	1	0	0	1	1	4
11	1	0	1	0	0	3
12	1	0	1	1	0	1
13	1	1	0	0	1	1
14	1	1	0	1	1	2
15	1	1	1	0	0	1
16	1	1	1	1	0	2

*In the columns with causal or outcome conditions, the number 1 indicates the presence of a condition, or "yes"; 0 indicates its absence, or "no."
†The number of districts is reported simply to remind the reader that each row of a truth table may represent more than one case.

because they have the same outcome and differ on only one causal condition: the presence or absence of teachers' unions. This simplification strategy follows the logic of an experiment. Only one condition at a time is allowed to vary (the "experimental" condition). If varying this condition has no discernible impact on the outcome, it can be eliminated as a factor. Thus, the comparison of rows 13 and 14 results in a simpler combination:

TRACKING = RACE·CLASS·elections

This rule for combining rows of the truth table as a way of simplifying them can be stated formally: If two rows of a truth table differ on only one causal condition yet result in the same outcome, then the causal condition that distinguishes the two rows can be considered irrelevant and can be removed to create a simpler combination of causal conditions (a simpler term).

The process of combining rows to create simpler terms can be carried on until no more simplification is possible. Table 6.4 shows all the simplifications that are possible for the truth table in Table 6.3, using presence of ability tracking as the outcome of interest. In Table 6.4, the truth table rows from Table 6.3 with outcomes of "1" (presence of tracking) have been translated into the upper- and lowercase names in the manner just described. Panel A of this table simply lists the eight kinds of districts that track students according to ability. Panel B shows the first round of simplification. Each of the terms from panel A can be combined with one or more other terms to create simpler terms. Whenever two terms with four conditions are combined, the new term has three conditions because one condition has been eliminated.

Panel C of Table 6.4 shows the second round of simplification. In this round, terms with three conditions (from panel B) are combined to form terms with two conditions. For example, the term labeled #17 in panel B (race·class·UNIONS) can be combined with the term labeled #21 (race·CLASS·UNIONS) to form a two-condition term (race·UNIONS). All the terms from panel B combine with one or more terms from the same panel to produce the 3 two-condition terms listed in panel C.

The three terms in panel C can be represented in a single statement describing the conditions under which tracking in these suburban school districts occurs:

TRACKING = race·UNIONS + RACE·elections + elections·UNIONS

Tracking occurs when

1. Racial diversity is absent and teachers' unions are present;

2. Racial diversity is present and competitive school board elections are absent; or

3. Competitive school board elections are absent and teachers' unions are present.

Table 6.4   Simplification of Truth Table for Tracking (Table 6.3)

Panel A. Districts That Track Students

Rows	Causal Configurations
2	race·class·elections·UNIONS
4	race·class·ELECTIONS·UNIONS
6	race·CLASS·elections·UNIONS
8	race·CLASS·ELECTIONS·UNIONS
9	RACE·class·elections·unions
10	RACE·class·elections·UNIONS
13	RACE·CLASS·elections·unions
14	RACE·CLASS·elections·UNIONS

Panel B. First Round of Simplification

					Simplified Terms	Label for New Term
Rows	2	+	4	→	race·class·UNIONS	#17
Rows	2	+	6	→	race·elections·UNIONS	#18
Rows	2	+	10	→	class·elections·UNIONS	#19
Rows	4	+	8	→	race·ELECTIONS·UNIONS	#20
Rows	6	+	8	→	race·CLASS·UNIONS	#21
Rows	6	+	14	→	CLASS·elections·UNIONS	#22
Rows	9	+	10	→	RACE·class·elections	#23
Rows	9	+	13	→	RACE·elections·unions	#24
Rows	10	+	14	→	RACE·elections·UNIONS	#25
Rows	13	+	14	→	RACE·CLASS·elections	#26

Table 6.4   (Continued)

Panel C. Second Round of Simplification

				Simplified Terms
#17	+	#21	→	race·UNIONS
#18	+	#20	→	race·UNIONS
#18	+	#25	→	elections·UNIONS
#19	+	#22	→	elections·UNIONS
#23	+	#26	→	RACE·elections
#24	+	#25	→	RACE·elections

Before accepting these tentative results, it is important to determine if further simplification is possible, as is often the case. Sometimes the process of combining rows to produce simpler terms (presented in Table 6.4) generates "surplus" terms. A *surplus term* is redundant with other terms and is not needed in the statement describing the combinations of conditions linked to an outcome. In short, some of the terms that are left after the process of combining rows, just described, may be superfluous. Recall that the goal of comparative analysis is to describe diversity in a simple way. If the results can be further simplified by eliminating surplus terms, as is the case here, it is important to do so. The idea of surplus terms is best understood by examining the methods used to detect them.

The best way to check to see if there are surplus terms is to construct a chart showing which of the original terms in panel A of Table 6.4 are covered by which simplified terms in panel C of Table 6.4. A simplified term covers a truth table row if the row is a subset of the simplified term. For example, RACE·CLASS·elections·UNIONS (row 14 of the truth table) is a subset of the simplified term "elections·UNIONS."

The chart showing the coverage of the simplified terms is presented in Table 6.5. The simplified term "race·UNIONS" covers the first four terms from panel A of Table 6.4, while the term "RACE·elections" covers the other four. The third simplified term (elections·UNIONS) does not cover any of the rows uniquely; it covers two that are covered by the first simplified term and two that are covered by the second. Thus, the third simplified term is surplus; it is redundant with the other terms.

By eliminating the third simplified term, the results of the analysis of configurations can be reduced to

$$TRACKING = race \cdot UNIONS + RACE \cdot elections$$

This completes the procedure. The final statement says that tracking occurs

1. When racial diversity is absent and teachers' unions are present, or

2. When racial diversity is present and competitive school board elections are absent.

The first term (race·UNIONS) indicates that in school districts that are predominantly white, tracking is implemented if there are teachers' unions. This finding supports the researcher's belief that teachers' unions prefer tracking and specifies the conditions under which their interests are realized—in districts where there is an absence of racial diversity. It does not matter whether school board elections are open and competitive or whether the district contains a broad range of income groups. The second term (RACE·elections) indicates that in school districts where there is racial diversity, tracking occurs when school board elections are not competitive. They are routine matters that do not attract a lot of voter interest. In these districts, it does not matter whether teachers' unions are present or whether the district contains a broad range of income groups. The second term suggests that if voters become involved in school board elections, tracking would be eliminated in racially diverse districts.

The analysis of school districts presented here shows the major steps in using comparative techniques to unravel causal patterns:

1. Select causal and outcome conditions, using existing social science literature and substantive knowledge to guide the selection.

2. Construct a sorted data table showing the scores of cases on these causal and outcome conditions (Table 6.2).

3. Construct a truth table from the data table, making sure that cases with the same causal conditions actually have the same score on the outcome (Table 6.3).

4. Compare rows of the truth table and simplify them, eliminating one condition at a time from pairs of rows (Table 6.4).

5. Examine the coverage of the simplified terms to see if there are any surplus terms that can be eliminated (Table 6.5).

Table 6.5   Chart Showing Coverage of Simplified Terms

Truth Table Rows*	Simplified Terms†		
	race·UNIONS	RACE·elections	elections·UNIONS
race·class·elections·UNIONS	✓		✓
race·class·ELECTIONS·UNIONS	✓		
race·CLASS·elections·UNIONS	✓		✓
race·CLASS·ELECTIONS·UNIONS	✓		
RACE·class·elections·unions		✓	✓
RACE·class·elections·UNIONS		✓	
RACE·CLASS·elections·unions		✓	
RACE·CLASS·elections·UNIONS		✓	✓

*From panel A of Table 6.4.
† From panel C of Table 6.4.

The terms that remain after step 5 show the simplest way to represent the patterns of diversity in the data. In the comparative analysis presented in Tables 6.2 through 6.5, the goal was to explain why some school districts track elementary students. The results show which types of school district track elementary students and distinguish them from those that do not.

Conclusion

The brief overview of comparative methods presented in this chapter illustrates some of the key features of the comparative approach. The most important feature is its focus on diversity. Whenever a set of cases has different outcomes (cities with different reactions to Somali refugees, countries with different reactions to IMF-mandated austerity programs, bars with different ways of encouraging patrons to drink responsibly, and so on), comparative methods can be used to find simple ways of representing the patterns of diversity that exist among the cases. These methods identify similarities within subsets of cases that distinguish them from other subsets.

As in all forms of social research, analytic frames and images play an important part in comparative research. Analytic frames provide primary leads for the construction of truth tables, especially the selection of causal conditions. The construction of the truth table itself is an important part of the dialogue of ideas and evidence in comparative research because the truth table must be free of inconsistencies before it can be simplified. Evidence-based images emerge from the simplification of truth tables in the form of configurations of conditions that differentiate subsets of cases.

In many ways, the comparative approach lies halfway between the qualitative approach and the quantitative approach. The qualitative approach seeks in-depth knowledge of a relatively small number of cases. When the focus is on commonalities, it often narrows its scope to smaller sets of cases as it seeks to clarify their similarities. The comparative approach usually addresses more cases because of its emphasis on diversity, and it is applied to sets of cases that are clearly bounded in time and space. By contrast, as Chapter 7 will show, the quantitative study of covariation seeks broad familiarity with a large number of cases and often views them as generic, interchangeable observations.