CHARLES C. RAGIN

FUZZY-SET SOCIAL SCIENCE

THE UNIVERSITY OF CHICAGO PRESS
CHICAGO AND LONDON

Charles C. Ragin holds a joint appointment as professor of sociology and political science at Northwestern University. He also holds an appointment in the Department of Sociology and Human Geography at the University of Oslo, Norway. He has published over fifty journal articles and book chapters. His books include *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*, *Issues and Alternatives in Comparative Social Research*, and *Constructing Social Research: The Unity and Diversity of Method*.

The University of Chicago Press, Chicago 60637
The University of Chicago Press, Ltd., London

Printed in the United States of America
09 08 07 06 05 04 03 02 01 00 1 2 3 4 5
ISBN: 0-226-70276-6 (cloth)
ISBN: 0-226-70277-4 (paper)

Library of Congress Cataloging-in-Publication Data

Ragin, Charles C.
Fuzzy-set social science / Charles C. Ragin.
p. cm.
Includes bibliographical references and index.
ISBN 0-226-70276-6 (alk. paper) — ISBN 0-226-70277-4 (pbk. : alk. paper)
1. Social sciences—Research. 2. Fuzzy sets. I. Title.
H62.R233 2000
300'.1'511322—dc21 99-086553

∞ The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48–1992.

CONTENTS

ILLUSTRATIONS

THREE

STUDYING CASES AS CONFIGURATIONS

. . . the meaning of a low export ratio for Bohemia, a small country surrounded by the rest of the European world-economy, and a similar low ratio for Russia, a large empire on the edge of the European world-economy, must have been quite different. Bohemia's freedom of political action was ultimately far smaller and hence her economic dependence ultimately far greater.

—Wallerstein, *The Modern World-System*

INTRODUCTION

Social scientists who use conventional variable-oriented methods are not especially fond of arguments like the one just quoted. Essentially, Wallerstein is arguing that even though two cases (Bohemia and Russia in the sixteenth century) had roughly the same scores on a key variable (value of exports relative to the value of total domestic production—a key indicator of trade dependence), the interpretation of the data is different. The meaning of these scores is shaped by context—by other features of the cases in which these two, more or less identical scores are embedded. From the perspective of variable-oriented social science, the idea that a score must be interpreted *in context* essentially argues against the common view that aspects of cases can be evaluated separately from each other, especially with respect to their "independent" causal effects on some outcome. Indeed, the logic of conventional variable-oriented social science is explicitly organized around isolating the effect of each causal variable—estimating its effect on some outcome, net of the effects of competing variables.

Imagine a survey researcher arguing that an age of "65 or older" means very different things for different people, depending on other

relevant features, among them gender, race, religion, and social class. This position would put an end to the common practice of treating age as an "independent" variable and estimating its separate or "unique" impact on relevant outcomes—for example, its net impact on voting preferences, net of the effects of other relevant causal variables. Following the line of reasoning indicated in the epigram at the beginning of this chapter, we would expect to find that the impact of age on voting preferences depends on context—on other relevant aspects of each voter. For some individuals being "65 or older" might make them more conservative; it might make others more liberal; it might make others more apathetic; it might have no impact on others; and so on.

Conventional variable-oriented research usually does not pay close attention to the context of a score on a variable. Consider the usual approach to measurement: First, the researcher identifies important features of cases—key dimensions of cross-case variation. A researcher studying the world economy, for example, might identify as one such aspect the degree to which different countries participate in international trade. Next, the researcher develops a measure or indicator of this feature and then proceeds to derive a score for each case. In this example, the researcher might use a common indicator such as the ratio of the value of exports to the value of gross domestic product (GDP) for each country. After obtaining scores for all relevant countries, the investigator then computes an average score and uses this measure of central tendency to define which countries have "low" scores and which countries have "high" scores. In this approach, countries well above the mean have "high" scores; countries well below the mean have "low" scores.

The notion of assessing scores *in context*, indicated in the epigram, challenges the conventional practice of equating cases with similar scores on a given variable. According to this alternate reasoning, two cases can have identically "high" or identically "low" scores, but without looking at each score's context, it would be very hazardous to equate them.[1] At its core, this alternate view questions an implicit assumption

1. Actually, most conventional variable-oriented methods are even less anchored in context than my discussion indicates. These methods not only equate similar scores, they also equate similar differences between scores. For example, a $1,000 difference in GNP per capita is treated as the same regardless of where it occurs in the range of GNP per capita values. Thus, the differences between $500 and $1500 in GNP per capita is equated, in causal impact, to the difference between $14,000 and $15,000. Anyone who has toured both less developed and more developed countries knows that the first difference is dramatic; the second is trivial.

of the variable-oriented approach: that the cases included in a sample are homogeneous enough to permit equating similar scores. Looking at each score's context provides a way to evaluate the assumption of homogeneity that is often entailed in the constitution of conventional populations and samples (e.g., the common practice of treating a set of "countries" as a sample). Returning to the epigram, it is clear that Wallerstein is arguing that it is hazardous to equate these two countries' low scores precisely because the two cases belong to different populations of countries. That is, considering their contrasting size and geographic location, they can be seen as different kinds of cases.

The principle that "context matters" is central to the configurational approach to cases. As noted in chapter 1, one of the key features of case-oriented research is its attention to cases as configurations of aspects, conceived as set memberships. As I show in this chapter, a configurational understanding of cases is central to diversity-oriented research. To view each case as a configuration, it is necessary to examine relevant aspects of a case all at once, as an interpretable combination of elements.

I develop this idea of cases as configurations first by presenting an overview of the understanding of cases that is integral to the single-case study. This research strategy typically refrains from explicit cross-case comparisons and instead emphasizes how aspects of a single case interconnect. I then extend the idea of configurational analysis to the study of multiple cases and show the link between looking at cases as configurations and studying their diversity. The key to understanding cases as configurations is to view them in terms of the different *combinations* of relevant attributes they exhibit. By grouping cases into a relatively small number of configurations of attributes, the researcher establishes a basis for specifying different "kinds" of cases. In this way, the researcher can understand types of cases as different configurations of attributes.

In the second half of this chapter I show how the idea of cases as configurations resonates with the "property space" approach to typology construction first advocated by Paul Lazarsfeld in the 1930s. In essence, Lazarsfeld (1937), and later Allen Barton (1955), advocated a configurational approach to the problem of reducing the complexity of social phenomena (see also Lazarsfeld and Rosenberg 1955; Lazarsfeld et al. 1972). Their approach, in effect, allows a researcher to translate a multidimensional attribute space into a handful of types. I extend their approach by showing that it is possible to formalize their technique of "functional reduction." This formalization, in turn, provides

important tools for interrogating property spaces, revealing the ways in which the diversity within a given property space may be "limited."

This chapter also provides a demonstration of the use of configurations as analytic units, an approach that is integral to diversity-oriented research. While I limit my focus to the use of presence/absence dichotomies in this chapter (and in all of part 1), in part 2 I demonstrate how to extend these principles to phenomena that also vary by level or degree, using fuzzy sets.

THE LOGIC OF CASE-STUDY RESEARCH

Many social scientists conduct case studies. This research strategy, in fact, may be the most common form of social scientific inquiry. It is by far the most popular research strategy in anthropology and history, and it is much more common in political science and sociology than most scholars realize (Feagin, Orum, and Sjoberg 1991). Many researchers, however, are reluctant to see their work as case-study research. After all, one of the central goals of social science is to generalize, and social scientists are trained to be wary of drawing general conclusions from a single case. What's one case? The empirical world displays a great deal of randomness and unpredictability, according to this reasoning. General lessons drawn from a single case, therefore, must be inherently suspect (Sjoberg et al. 1991).

Many studies of cross-case patterns appear to be based exclusively on the analysis of large *N*s when in fact they are also case studies. A study of the changing relationship between income and single parenthood in the United States over the post–World War II period, for example, can be seen as a case study, even though it might involve quantitative analysis of census and survey data on thousands of households, each conceived as a separate "case" in the quantitative analysis. In the end, the study is also about the United States in the second half of the twentieth century, not just the many individuals and families included in the analysis.[2] More than likely, the explanation of the changing relation between income and single parenthood would focus on interrelated aspects of the United States over this period. For example, to explain the weakening link between low income and single parenthood the researcher might cite the changing status of women, the decline

2. Michael Burawoy (in Burawoy et al. 1991) has developed this expansive notion of case-study research in his concept of the "extended" case method. His own work (e.g., Burawoy 1979) exemplifies this strategy.

in the social significance of conventional family forms, the increase in divorce, the decrease in men's job security, and other changes occurring in the United States over this period.

The logic of the case study is fundamentally configurational. Different parts of the whole are understood in relation to one another and in terms of the total picture or package that they form. The central goal is usually to show how different "parts" of a case interconnect, for example, how a weakening of the link between low income and single parenthood connects with other changes in the United States over the second half of the twentieth century. The "parts" that case-study researchers examine can be quite varied: institutions, path dependencies, social structures, historical patterns and trends, routine practices, singular events, event sequences, connections to other cases, the case's larger environment, and so on. While case-study researchers are sometimes tempted to take this approach an additional step and argue that the case in question involves a unique configuration of parts and that it could be constituted in one and only one way, this additional step is usually unwarranted. What matters most is that the investigator makes sense of multiple aspects of the case in an encompassing manner, using his or her theory as a guide.

Donald Campbell (1975) offers a rough formalization of this approach in his ruminations on the logic of case-study research. At first glance, Campbell argues, the case study appears to be totally lacking in scientific merit because there is only one case to explain and many possible explanations to choose from. He notes, however, that case-study researchers often reject theories because they do not explain the facts of their case. Further, despite having only one case, they often must struggle to find theories that work (see also Walton 1992). Why is it so difficult? The key to this puzzle is the simple fact that every theory has many implications, relevant both to features of the case in question and to causal processes and sequences operating within the case. Thus, the case-study researcher evaluates many theoretical implications relevant to his or her case to see if the case conforms to expectations (King et al. 1994). Not all features of the case are compatible with the initial theory, and the case-study researcher must either find an alternate theory that works better, revise an existing theory, or propose an entirely new one (Walton 1992).

Campbell (1975) suggests that each separate theoretical implication can be seen as a separate "observation" for "testing" a theory. Thus, a single case becomes many observations—some contradicting and some supporting competing theories. Collectively, competing theories and their different implications define all theoretically relevant aspects of the case in question. The researcher's task is to see which theory does the best job of explaining aspects of the case relevant not only to its own implications but also to the implications of competing theories. Thus, after defining and selecting relevant aspects of the case, the case-study researcher assesses the explanatory power of each theory. The theory (or combination of compatible theories) that best covers both its own implications and those of competing theories prevails and provides the basis for the investigator's representation of the case.

In the end, the researcher crafts an explanation, embedded in his or her representation of the case, that satisfies as many theoretical implications as possible in a coherent manner. The success of a case study hinges on (1) the number of relevant aspects of the case the researcher can encompass with his or her explanation, (2) the success of the researcher in showing that his or her portrait of the case actually makes sense of all the aspects that he or she has deemed theoretically relevant, and (3) the agreement of other scholars that all relevant aspects of the case in question have, in fact, been addressed by the researcher in a convincing manner. Of course, the most successful case studies accomplish much more than theoretical and substantive coherence. They also may advance theory (Burawoy et al. 1991) or establish important lessons for policymakers. Still, theoretical and substantive coherence should be considered preconditions for these more ambitious goals (Yin 1994).

COHERENCE AND CONFIGURATIONS IN CASE-STUDY RESEARCH

Whenever a researcher investigates a case with an eye toward how the different parts or aspects of a case interconnect, the inquiry has a configurational character. The organizing idea in such research is that the parts of a case constitute a coherent whole—that they have an integrity and coherence considered together. For example, researchers who study family systems often find that patterns of interpersonal accommodation are so enmeshed that a "dysfunction" cannot be remedied without addressing many different aspects of the family all at once. Likewise, researchers who study cultures often observe that cultural traits come in packages that seem to defy disassembly. Such configurations do not defy analysis, per se, because they can be viewed in terms of interrelated parts or aspects.

Campbell's (1975) argument that case-study researchers test multiple theoretical implications underscores the configurational character of this type of research. Because theoretical implications direct the case-study researcher's attention to "observations" (i.e., different aspects of a single case) that cannot be independent of one another, the researcher must make sense of them all at once, as a package. Furthermore, because different theories typically have implications about different aspects of the case, the researcher's attention is directed to a broad range of aspects, all of which may be connected in some way. Thus, case-study researchers examine overlapping configurations of aspects as they weigh the relative explanatory power of competing theoretical perspectives.

Because of their configurational nature, case studies often have a house-of-cards quality. A configurational understanding of a case can fall apart all at once if contradictory evidence is introduced. A central principle of configurational thinking that is relevant to social scientific inquiry is the idea that the character of the "whole case" may change qualitatively if a single key part is altered or changed in some way. For example, the conventional portrait of the Holocaust during World War II as the obsession of Hitler and his inner circle crumbled for many scholars with the publication of Daniel Goldhagen's *Hitler's Willing Executioners: Ordinary Germans and the Holocaust* (1996). The conventional account could not assimilate Goldhagen's new evidence, and the whole portrait changed.

It follows from these observations on case-study research that this research strategy is extraordinarily theory-dependent. Theory is used to make sense of the case as a configuration of theoretically relevant aspects, which cohere as a package. This package, in turn, can be remarkably fragile: the strands of theory that bind it together can easily unravel. Because case studies are configurational, a change in the evidence on a single key aspect of a case can change the character of the whole qualitatively. In this light, the common charge that case-study research is merely descriptive appears ludicrous. While many case-study researchers keep their theories hidden or implicit, the entire enterprise is sheer chaos without some form of theoretical guidance.[3]

3. It is also possible to view the proposition that "the case is a coherent configuration" as a hypothesis that can be evaluated and refuted.

MULTIPLE CASES, MULTIPLE CONFIGURATIONS

The principles of configurational thinking just described are relevant to the study of social phenomena in general, not just to the study of single cases. However, most social scientists who conduct cross-case analyses adhere to the variable-oriented approach, which in turn tends to treat cases not as configurations of aspects but as collections of distinct, analytically separable attributes. In the conventional variable-oriented view, to change one aspect of a case results in a case that differs only slightly from the original, not one that may differ qualitatively. In short, the conventional variable-oriented approach eschews configurational thinking.

Consider, for example, the study of individual-level differences. In conventional variable-oriented research, cases are viewed as "different" when they differ on many aspects and "similar" when they differ on only a few. For example, two cases with mostly the same values on predictor variables receive roughly the same predicted values on the outcome variable in multiple regression analysis.[4] In this framework, similarity and difference can be assessed in a straightforward accounting manner. For example, a White, suburban, middle-class, male professional with three children and a mortgage who votes for the Republican Party differs by only a single trait from a person with all these same characteristics but who votes for the Socialist Labor Party. Thus, from a variable-oriented viewpoint, they are very similar. Still, it might be very different to be trapped in an elevator with one versus the other. This qualitative difference captures the essence of viewing cases as configurations: Two cases may be similar in most ways but because they differ on one or more key aspects, their difference may be one of kind, not simply one of degree.[5] Furthermore, if two cases do differ qualitatively, it is hazardous to equate their similarities. These configurational notions are foreign to variable-oriented social science, at least as it is conventionally practiced.

The de facto repudiation of configurational thinking is integral to variable-oriented research strategy. It is clearly evident in this strategy's

4. The multiple regression framework also allows "compensation," which makes it possible for cases that differ in counterbalancing ways to have the similar predicted values on the outcome variable.

5. It is important to emphasize that in the configurational approach a single difference has the potential to warrant a qualitative distinction. It is not the case, however, that every difference justifies a qualitative distinction.

approach to the analysis of cross-case patterns. When confronted with more than a handful of cases, most researchers simply abandon the idea of cases as interpretable configurations and instead search for cross-case patterns. The search for cross-case patterns is usually correlational: researchers try to see if aspects of cases, conceived as variables, cor-relate. For example, a researcher might ask, ignoring Wallerstein's warnings, whether countries that are less trade-dependent have more political autonomy (i.e., from the countries that dominate the world economy) than countries that are more trade-dependent. This re-searcher would devise measures of both variables (trade dependence and political autonomy) and then assess their correlation across a range of countries to see if a strong inverse relation exists. Following Wal-lerstein's argument, however, one might object to this analysis because it ignores the different contexts of countries' scores on these variables and thus may equate scores that should not be equated.

The problem just sketched is *not* that the analysis, as described, is merely bivariate. In other words, the ignorance of context could not be remedied by controlling for the effects of country size and location on autonomy in a multivariate analysis of the impact of trade depen-dence on autonomy. To control for the effects of country size and loca-tion on autonomy does not address the core issue because the basic idea behind configurational thinking is that aspects of cases should be examined together, as packages.[6] By viewing aspects of cases configu-rationally, it is possible to assess whether the impact of similar scores (e.g., degree of trade dependence) on some outcome (e.g., political au-tonomy) differs by context.

It is important to detail, at this juncture, how a configurational ap-proach differs from a conventional variable-oriented approach in its conception of cases and variables. In its most basic form, the idea of viewing cases as configurations can be captured by examining different combinations of values on relevant variables and treating each combi-nation of values as a potentially different type of case. Consider the following illustration, which builds on the discussion of export depen-dence, country size, and country location. To keep the illustration sim-

6. For readers familiar with the language of statistical analysis, my argument here is that to study cases as configurations, it is necessary to start with a model of saturated interaction, using all relevant variables, and then to simplify this model in a top-down manner. In most social research, such models are very diffi-cult to estimate, and when they can be estimated, they are very difficult to interpret. Typically, the interaction terms are highly collinear with each other and many different interaction models may fit a given data set equally well.

Table 3.1
From Variables to Types

Country Type	Variable		
	EXPORT RATIO	SIZE	LOCATION
A	Low	Small	Close
B	Low	Small	Far
C	Low	Large	Close
D	Low	Large	Far
E	High	Small	Close
F	High	Small	Far
G	High	Large	Close
H	High	Large	Far

ple, assume that the main causally relevant aspects of interest (export ratio, country size, and location) are all dichotomies. Thus, export ratio is either "high" or "low;" size is either "small" or "large"; and location is either "close to" or "far from" the core of the world economy. With three dichotomies, there are 2^3 (i.e., 8) logically possible combinations of values. The key to configurational thinking is to see these eight combinations of values as providing the basis for differentiating eight different kinds of cases, which, in turn, may constitute eight different "populations." In other words, from a configurational perspective the analysis of country characteristics just described involves not three independent variables, but eight configurations conceived as types of cases. The transformation of these three variables into eight types of cases is shown in table 3.1.

In this alternate conception of the relevant analytic space, Wallerstein's Bohemia is type A and his Russia is type D. That is, these two countries are treated as distinct kinds of countries, not as two countries that have similarly low export ratios, while differing in size and location. In this scheme a type-A country (low export ratio, small size, close to core) may be as different from a type-B country (low export ratio, small size, far from core) as it is from a type-H country (high export ratio, large size, far from core), even though a type-A country differs by only one attribute from a type-B country, while it differs by three attributes from a type-H country.

This view of cases as configurations implements the key principle of configurational thinking previously sketched—the idea that a single difference between two cases may constitute a difference in kind. Table 3.1 shows that by delineating all possible combinations of values of the relevant variables it is possible to specify different configurations, which in turn establishes a framework for delineating types of cases.[7] Thus to study cases as configurations it is not necessary to abandon variables altogether. Rather, it is necessary simply to abandon the idea that variables should be seen as independent, separable aspects of cases. Instead, variables should be seen as the components of configurations.

THE LINK BETWEEN STUDYING CONFIGURATIONS AND STUDYING DIVERSITY

While the configurational understanding of cases has its roots in case-oriented research and the study of small Ns, it is integral as well to diversity-oriented research. As explained in chapter 1, diversity-oriented research lies midway between studying general patterns across "all" relevant cases (conceived as homogeneous members of a single population), on the one hand, and attending to the complexity of a specific case or a narrowly circumscribed set of cases, on the other. To study diversity is to allow for the possibility that the cases included in a study differ by type or kind. As just demonstrated, the study of cases as configurations provides a useful framework for the specification of types. In this approach, conventional variables are conceived as raw material for delineating types, not as analytically distinct, independent attributes.

In diversity-oriented research, investigators are ever-conscious of the possibility that similarities among cases may be illusory. At first glance, the cases included in a study may seem similar, as different instances of the same general phenomenon. In the course of the research, however, focusing on configurational differences among cases may challenge the initial perception that the cases are all of the same kind. Instead, the researcher may differentiate types of cases. Attending to diversity thus involves careful consideration of the possibility that cases differ by type or kind, not merely by level or degree.

For example, consider a researcher studying gender and work in twenty microelectronic assembly plants in southeast Asia.[8] The initial impression might be that women are exploited everywhere, with modest variation in the degree or intensity of their exploitation. Close examination of these twenty cases with an eye toward their diversity,

7. As I show in part 2, this approach is not limited to dichotomies. Using fuzzy sets, it is possible to assess the degree to which each case conforms to each configuration and to use this information to evaluate causal conditions.

8. This example draws inspiration from the work of Leslie Salzinger (1997).

however, might challenge this simplistic conclusion. For example, the twenty plants might differ in how gender is intertwined with the labor process. In some factories, anarchy on the shop floor may permit men to dominate women and keep them in subordinate work roles. In others, management may hire women for production tasks and men for supervisory roles and then consciously manipulate traditional gender roles to motivate women to work harder. Examination of the remaining plants might reveal other distinct ways of manipulating gender in the workplace. All twenty might still be seen as instances of gender exploitation, and one could arguably rank them according to degree of gender exploitation. But the key finding in a study emphasizing diversity would be the different ways of gendering work exhibited by these cases. The conclusion would be that different ways of organizing work—conceived as configurations of hiring, shop floor, and other practices—give rise to different forms of gender exploitation.

In practical terms, this study would proceed simultaneously as a set of case studies and as a search for clusters of commonalities across subsets of cases. The researcher would use available theories to identify theoretically important aspects of assembly plants, especially those features relevant to the gendering of work. Examination of specific cases would help the researcher pinpoint the most important features of cases and prod him or her to consider how they might constitute coherent packages (i.e., as interconnected practices and structures). If two or more plants seemed similar in how they gender work, the researcher would try to identify the core elements that structure work in similar plants and establish preliminary types. Close examination of cases would also permit an assessment of the adequacy of current theories, which in turn would lead the researcher either to press on with his or her initial theoretical formulation or adopt a new one. After achieving some confidence that the important features of cases had been identified, the researcher could then specify different configurations of these aspects, looking at their combinations (as in table 3.1), and thus establish and refine the conceptual basis for delineating types of cases.

The important point here is that in the initial stage of a project, the researcher focuses on key aspects of cases relevant to some outcome, such as gender exploitation. Once identified, the researcher examines these aspects in various combinations to see which aspects form coherent packages and bring together cases that seem the same, while distinguishing cases that seem different. Thus, the researcher focuses on specific features of cases, identifying the most relevant, while at the same time considering how these features cohere with each other as distinct packages, especially across cases. The end result is a delineation of types, based on specific empirical configurations. Thus, examining cases as combinations of aspects provides the primary means for mapping their diversity. Indeed, looking at cases as configurations and allowing for the possibility that a single difference may constitute a difference in kind highlights their diversity.

It also should be noted that in the process of identifying the different configurations of attributes that make up types, the researcher also specifies, implicitly and by default, the combinations of features that do not exist, based on the evidence examined in the study. Some of these combinations may be infrequent or unlikely; others may involve combinations of characteristics that are impossible, incompatible, or simply unwieldy. In diversity-oriented research it is important to consider which combinations are not found and to explore the possible reasons for their absence. For example, if the researcher in the example just described found that there were no microelectronic assembly plants that combined internal labor markets and the conscious manipulation of traditional gender roles, it would be worthwhile to consider the reasons for the apparent incompatibility of these two features.

CONFIGURATIONS, TYPES, AND LAZARSFELD'S CONCEPT OF PROPERTY SPACE

One reason that the study of diversity, as outlined in this work, has not received appropriate attention over the past several decades is the simple fact that social scientists have conceived the problem of typology construction almost exclusively from the perspective of variable-oriented analysis. When variable-oriented researchers discuss types, they usually focus on statistical procedures such as cluster analysis and Q-factor analysis (Bailey 1994). These techniques typically use as input a data matrix of cases by variables and apply various computer algorithms to sort cases into a small number of groups. These procedures focus on similarities and differences among cases across a range of variables specified by the investigator. The usual goal is to inductively derive a small number of groupings such that within-group differences are minimized while between-group differences are maximized.

Typically, the number of groups that "emerge" from the application of these procedure is arbitrary, and it can vary widely from one analysis to the next, depending on the specific variables included, how they are measured, the nature of the clustering algorithm used, the way in which similarity is calculated, the researcher's tolerance for complexity, and so on. Unfortunately, the procedures embedded in the computer algorithms that induce types are invisible to most researchers, and few researchers have the experience or expertise to differentiate between findings and methodological artifacts. Because the results of these procedures can vary so greatly from one analysis to the next, researchers often become frustrated and simply abandon the effort to formulate types altogether. Besides, researchers are not always sure what to do with a typology once it is derived. Should it be conceived as a categorical variable in a multivariate analysis, or should it instead provide a basis for differentiating qualitatively distinct populations or subpopulations? And how should researchers treat cases that vary in how well they conform to types?

Modern methods of typology construction such as cluster analysis result in classification schemes that are almost always "polythetic" and usually "fully polythetic" (Bailey 1994). In a polythetic classification, the cases that are grouped into a single category are allowed to differ substantially from each other on one or more attributes as long as they are similar on many of the attributes selected by the investigator for inclusion in the analysis. In a fully polythetic classification, the cases grouped into a category may fail to exhibit a single commonality. Thus, for example, a fully polythetic grouping might include individuals who are mostly but not all Republican, mostly but not all middle class, mostly but not all White, mostly but not all suburban, and so on.

Clearly, polythetic schemes, and especially fully polythetic schemes, violate the core principle of configurational thinking previously sketched, namely, the idea that a single difference may constitute a qualitative distinction or a difference in kind. Recall the stuck elevator example: Two individuals are similar in all respects save one. Despite their many similarities, it might be very different to be stuck on an elevator with one versus the other. Using modern methods of cluster analysis, however, these two individuals would almost certainly be assigned to the same cluster. From a configurational perspective, by contrast, they might be seen as qualitatively different and thus assigned to different types.

The configurational principles advanced in this chapter hark back to a much earlier tradition of typology construction, the "property-space" approach developed by Lazarsfeld (1937) and elaborated by Barton (1955). Lazarsfeld argued that most "type concepts" involve sets of attributes that make sense together as a unitary construct, for example, Weber's (1978) specification of the ideal-typic bureaucracy. He noted further that too often social scientists use type concepts without analyzing them—that is, without examining their component attributes. He argued that it is important to identify the component attributes because they provide the basis for elaborating a full typology, based on the different logically possible combinations of attributes that make up a type concept. In this approach, the attributes of a type concept constitute a "property space" with as many dimensions as attributes. Each combination of attributes is a specific location in the property space. Each of these locations, in turn, may constitute a different type.

Suppose, for example, a researcher determines that the "authoritarian personality," a type concept, has four component attributes. A full typology can be constructed using these four components as presence/absence dichotomies to produce a four-dimensional property space, which, in this example, could be represented as a two-by-two-by-two-by-two table. Each cell of this four-way table is one of the logically possible combinations of the four dichotomies. In this property space, Lazarsfeld would argue, the authoritarian personality can be seen as one of 16 personality types formed from the four attributes. (With four presence/absence dichotomies there are 2^4 or 16 logically possible combinations.) In Lazarsfeld and Barton's approach, the four dichotomies and the sixteen combinations they yield constitute the "property space" that is implicit in the conception of the authoritarian personality, a single type concept. Their combinatorial approach to property spaces dovetails with the understanding of types depicted in table 3.1 and with the configurational principle that a single difference between two cases may constitute a difference in kind.

"REDUCING" PROPERTY SPACES

As Lazarsfeld and Barton both note, it is quite likely that the researcher will have difficulty finding empirical instances of all logically possible combinations of attributes. In the authoritarian personality example, for instance, it might prove impossible to find individuals lacking all four attributes. Likewise, in the example elaborated in table 3.1, it might prove impossible to find one or more of the eight types of countries. Lazarsfeld and Barton argue that the fact that many of the combinations in a property space lack empirical instances is a great boon to social research, especially when the number of component attributes, and thus the number of combinations, is great. The basic idea here is that a simple and useful empirical typology can be formed from the

combinations that exist. If a relatively small number of combinations exists empirically, then the researcher will be able to reduce a multidimensional property space to a handful of categories. Lazarsfeld (1937: 127–8) calls this simplification a *functional reduction*. The idea of functional reduction is echoed in Arthur Stinchcombe's remarks on typologies. He states that "a typology is a statement that a large number of variables have only a small number of combinations of values which actually occur, with all other combinations being rare or nonexistent. This results in a radical improvement in social scientific theory" (1968:47).

Consider as an example of this type of simplification (functional reduction of a property space) the hypothetical evidence on countries presented in table 3.2. This table borrows the three dichotomies used in table 3.1 (high versus low export ratio, large versus small size, close to versus far from the core of the world economy) and adds a fourth (exporter of raw materials versus exporter of finished goods). Imagine that the researcher in this example started with a type concept, "core country," and determined that this construct embraced four key attributes—large size, close proximity to the core of the world economy, high export ratio, and participation in the world economy as an ex-

Table 3.2
Example of Lazarsfeld's "Functional Reduction"

Combination No.	Variable				Substantial Empirical Instances?	Type
	NATURE OF EXPORTS	EXPORT RATIO	SIZE	LOCATION		
1	Raw	Low	Small	Close	No	
2	Raw	Low	Small	Far	Yes	4
3	Raw	Low	Large	Close	No	
4	Raw	Low	Large	Far	No	
5	Raw	High	Small	Close	No	
6	Raw	High	Small	Far	No	
7	Raw	High	Large	Close	No	
8	Raw	High	Large	Far	Yes	3
9	Finished	Low	Small	Close	No	
10	Finished	Low	Small	Far	No	
11	Finished	Low	Large	Close	No	
12	Finished	Low	Large	Far	No	
13	Finished	High	Small	Close	Yes	2
14	Finished	High	Small	Far	No	
15	Finished	High	Large	Close	Yes	1
16	Finished	High	Large	Far	No	

porter of finished goods. The elaboration of these four attributes as dichotomies yields the property space depicted in table 3.2. Suppose that the search for instances of the sixteen combinations, however, yielded substantial instances of only four of the logically possible combinations:

1. "core countries" (export finished goods, high export ratios, large in size, close to the core of the world economy);
2. "core-associated countries" (export finished goods, high export ratios, small in size, close to the core of the world economy);
3. "semiperipheral countries" (export raw materials, high export ratios, large in size, far from the core of the world economy); and
4. "peripheral countries" (export raw materials, low export ratios, small in size, far from the core of the world economy).

The researcher would conclude from this evidence that there are four types of countries, reflected in the four empirical combinations. The remaining twelve combinations, those lacking substantial empirical instances, are thus excluded from the typology. In the discussion that follows, combinations that lack empirical instances are called *hypothetical combinations*. They are hypothetical in the limited sense that they do not exist or are rare, not in the more abstract or theoretical sense that they all constitute "thought experiments."

Of course, the data that social scientists use are only rarely as tidy as that represented in table 3.2. Typically, when researchers construct property spaces, they find more empirical combinations than they could possibly anticipate. While not a central concern of this chapter, it should be noted that researchers can use various criteria to evaluate the frequency of combinations when deciding which are "empirically substantial" and which are "hypothetical." These criteria may be substantive, theoretical, probabilistic, or some combination of criteria. Substantive criteria, for example, might be used to define frequencies that constitute a critical mass. Theoretical criteria might be used to specify frequencies with properties of interest to a given theoretical principle or idea. Probabilistic criteria could be used to identify frequencies that exceed expected values given some model of randomness.

For example, a researcher might consider as empirically substantial only those combinations that reach a certain frequency threshold, defined using probabilistic criteria. With 16 logically possible combinations, the researcher might decide that the number of cases attached to a given combination must be significantly greater than 1/16 of the

total number of cases for the combination to qualify as empirically substantial (von Eye 1990). A standard one-tailed *z* test for the difference between proportions (Hayes 1973:213–4) or a binomial probability test could be used to evaluate any positive gap between the observed proportion of cases displaying a given combination and the benchmark proportion (1/16 or .0625 in this example). Combinations that meet this threshold would be considered empirically substantial; those that fall short would be considered hypothetical. Alternatively, the researcher might control for one or more of the marginal distributions of the attributes that make up the typology when making this assessment (von Eye 1990:59–142). The point here is simply that researchers may use a variety of formal criteria to evaluate the frequency of combinations, not that any specific technique for making these assessments is best.

Most property spaces contain many hypothetical combinations—configurations that either lack empirical instances altogether or have a substantively or statistically insignificant number of instances. In general, the larger the number of dimensions that comprise a property space, the greater the number of hypothetical combinations. Hypothetical combinations may exist for a variety of reasons. Sometimes they exist simply because the property space has more locations than there are relevant cases. After all, it takes only thirty-three dichotomies to generate more logical combinations than there are people on Earth. Alternatively, combinations may lack instances because they involve impossibilities—for example, pregnant males. Many other combinations lack instances because of the confounded nature of social phenomena. For example, if two attributes are strongly correlated, say race and neighborhood, then some combinations will be relatively rare (e.g., African Americans living in wealthy suburban communities). Because social phenomena tend to be highly confounded, researchers generally find an abundance of hypothetical combinations in multidimensional property spaces.

The observation that property spaces often have an abundance of hypothetical combinations underscores the simple fact that social phenomena typically exhibit "limited diversity" (Ragin 1987:104–13). Researchers confront limited diversity whenever they are unable to locate instances of all the types contained within their property spaces. Limited diversity stems from the confounded nature of social phenomena (e.g., overlapping inequalities, as just noted) and is exacerbated by the tendency for social characteristics to occur in syndrome-like clumps—to "harmonize." Almost all combinations of social attributes that occur with substantial frequency (e.g., families that are middle class, White, gun-owning, and politically conservative, with children enrolled in fundamentalist Christian private schools) are suggestive of syndrome-like patterns extending well beyond the listed attributes. For example, this clump of characteristics might extend to NRA membership, rural background, and so on—to complete a common stereotype. Because harmonizing diversity is common in social life, it contributes to the abundance of hypothetical combinations in property spaces.

USING SET THEORY TO DESCRIBE LIMITED DIVERSITY

It is important when examining the distribution of cases within a property space to consider the ways in which diversity is limited, to see if the voids in the property space are patterned and if so, how. An understanding of the limits on diversity provides insights into the nature of the phenomenon under investigation and useful leads for constructing types. When there is a strong pattern to the combinations lacking empirical instances, the investigator should consider the possible social bases for these voids and incompatibilities. For example, it is certainly not coincidental that there are no African American females in charge of any of the largest U.S. corporations. The fact that there are no instances of this combination says something about the United States and the culture of its major corporations.

Set theory can be used to describe, in a very concise manner, the limitations on the diversity of cases within a property space. This approach offers simple procedures for summarizing in a single statement the nature of the hypothetical combinations (i.e., those lacking empirical instances) found in a property space. For illustration consider table 3.3, which shows a property space for the attributes of the members of a social science department at a large state university. How can the limited diversity of characteristics within this property space be described in a concise manner?

The table lays out a four-dimensional property space using four dichotomies relevant to the diversity within the department in question: gender (male versus female), race/ethnic background (European American versus minority), rank (tenured versus not tenured), and pay (above versus below average, compared with faculty members with the same number of years in seniority). At first glance, it appears that the members of this department are quite diverse with respect to the four dichotomies: There are male and female members; members from different race/ethnic backgrounds; members at different ranks; and members who are paid above average and below average, relative to their years in seniority. However, as close inspection of table 3.3 reveals, not all sixteen logically possible combinations of these four dichotomies exist in the department; only nine do. Thus, there are seven hypothetical combinations of characteristics. How are these hypothetical combinations patterned?

The principles of set theory that provide the basis for translating these seven combinations lacking empirical instances into a single statement about limited diversity are simple and straightforward. In essence, set theory is used to group hypothetical combinations together, forming larger sets. When deriving these larger sets, the basic goal is to specify a small number of broad groupings that embrace as many hypothetical combinations as possible. For example, the two hypothetical combinations "male-minority-tenured-above" and "male-minority-tenured-below" (nos. 3 and 4 of table 3.3) can be joined to form the more inclusive combination, "male-minority-tenured." The greatest shorthand is achieved by deriving the smallest possible number of larger groupings with the smallest possible number of commonalities shared within each larger grouping. The end result is a reduction of the hypothetical combinations in a property space to a concise statement describing the limits on diversity that exist within it.

Table 3.3
Using Set Theory to Describe Hypothetical Combinations

Combination No.	Aspect				Empirical Instances
	GENDER	ETHNICITY	RANK	PAY	
1	Male	Minority	Untenured	Below	Yes
2	Male	Minority	Untenured	Above	Yes
3	Male	Minority	Tenured	Below	No
4	Male	Minority	Tenured	Above	No
5	Male	European	Untenured	Below	Yes
6	Male	European	Untenured	Above	Yes
7	Male	European	Tenured	Below	Yes
8	Male	European	Tenured	Above	Yes
9	Female	Minority	Untenured	Below	Yes
10	Female	Minority	Untenured	Above	No
11	Female	Minority	Tenured	Below	No
12	Female	Minority	Tenured	Above	No
13	Female	European	Untenured	Below	Yes
14	Female	European	Untenured	Above	No
15	Female	European	Tenured	Below	Yes
16	Female	European	Tenured	Above	No

The details of these analytic procedures are presented in *The Comparative Method* (Ragin 1987:85–124), where I describe what is known as the "Quine-McCluskey algorithm" for the reduction of *truth tables*, the set-theoretic terms for what social scientists call property spaces (McCluskey 1956; McDermott 1985; Mendelson 1970; Quine 1952; Roth 1975). These procedures are also implemented in the computer program *Qualitative Comparative Analysis* (Drass and Ragin 1992).[9] Applying these procedures to the hypothetical combinations in table 3.3 results in a simple logical statement with two large groupings:

hypothetical combinations = female·above + minority·tenured.

In this statement multiplication (denoted with midlevel dots) indicates combinations of characteristics (logical *and*); addition indicates alternate combinations (logical *or*). Translated to prose, the equation states simply that this social science department lacks two kinds of faculty members: (1) those who combine being female with being paid above average, and (2) those who combine being minority with being tenured. Inspection of table 3.3 reveals that the first grouping covers rows 10, 12, 14, and 16; the second covers rows 3, 4, 11, and 12. Thus, all seven hypothetical combinations are embraced by these two groupings.

The evidence on limited diversity indicates that neither minority and tenured nor female and better paid are compatible with this department. For whatever reason, individuals who display these two combinations are not found. Of course, more research on the department in question would have to be conducted in order for the researcher to attach substantive significance to these two voids. Perhaps the minority members are all relatively recent hires, with freshly minted Ph.D.'s. Perhaps the various contributions of the female faculty members to the department's teaching, service, and research missions are undervalued by those who set salaries. More than likely, the researcher studying this department would be offered many different and contradictory explanations for the two voids identified in the table.

Logical statements such as the one just derived that describe limitations on the diversity in a property space are useful because they provide a basis for interrogating and learning about property spaces. For example, it is possible to use the statement describing the hypothetical

9. QCA 3.0 (Drass and Ragin 1992) can be downloaded from the archive at <http://www.nwu.edu/sociology>. Fuzzy-Set/Qualitative Comparative Analysis [FS/QCA] 1.0 (Drass and Ragin 1999) offers all the procedures available in QCA along with procedures for analyzing fuzzy-set memberships.

combinations, just presented, to construct statements about the diversity of *empirically substantial* combinations. For example, because the statement shows that there are no minority members who are tenured, it is possible to deduce that (1) if a member is minority, he or she is also not tenured, and (2) if a member is tenured, he or she is not a member of a minority. These two deductions can be made because minority members are a subset of untenured faculty, and tenured faculty are a subset of nonminority members. Inspection of empirical combinations in table 3.3 confirms these two simple deductions. Similarly, because the statement shows that there are no female members with above-average pay, it is possible to deduce that (1) if a member is female, her pay is below average, and (2) if a member's pay is above average, the member is not female. In other words, females are a subset of members with below-average pay, and members with above-average pay are a subset of male members. Again, inspection of the empirical combinations listed in table 3.3 confirms these simple deductions. The patterns just identified also can be viewed in terms of overlapping divisions that are likely to fuel conflicts within the department, specifically, the intersection of being female and receiving below-average pay and the intersection of minority status and being untenured.

The analysis of the limitations on diversity in table 3.3 shows that even though the academic department appears to embrace considerable diversity, at least at first glance, its diversity is in fact strongly patterned. It is important to emphasize that the patterning of diversity demonstrated here is both revealed and highlighted by the configurational understanding of social phenomena. This view of cases provides the foundation for the construction of evidence as property spaces and the associated view of each row of a property space as a potentially different kind of case. As I have just demonstrated, the use of set theory extends the property space framework and allows the deduction of formal statements describing diversity, especially the limited diversity that exists within a property space.

LIMITED DIVERSITY AND THE CONSTRUCTION OF EMPIRICAL GENERALIZATIONS

There is another important reason to be concerned about limited diversity, beyond its resonance with viewing cases as configurations and the associated problem of constructing types. When diversity is limited, as it is in almost any study of naturally occurring social phenomena, researchers should always check to see if any of their conclusions mask "simplifying assumptions" about the evidence. When researchers con- struct generalizations using evidence that is limited in its diversity, they embed simplifying assumptions in their generalizations. This important principle is best understood by example.

Suppose, for instance, that a researcher decides to study the morale of the members of department depicted in table 3.3 and that, further, this researcher believes that the aspects listed in the table are the main causal conditions linked to morale differences. Based on several weeks of observation, the researcher concludes that the department's minority members are more demoralized than its nonminority members. In fact, there is a perfect pattern—the minority members are all demoralized, the nonminority members are not. This researcher draws the conclu- sion that minority status explains low morale.

In this hypothetical study, the investigator found a perfect relation- ship between minority status and morale and a weak relation between rank and morale (after all, minority members are only a subset of the nontenured). From this viewpoint, it seems quite reasonable to con- clude that minority status is all that matters as far as morale is con- cerned. But the investigator's conclusion is actually about members who are both minority and untenured, not simply about minority members, because there are no minority members who are tenured. To draw a conclusion about "the impact of minority status" without acknowledging the other relevant characteristic that minority faculty share is to assume implicitly that if there were any *tenured* minority members, then they would resemble the existing minority members in displaying the outcome in question (i.e., low morale).

This concern for simplifying assumptions inherent in the configura- tional view of cases contrasts sharply with conventional variable-ori- ented research where different aspects of cases tend to be examined as "independent variables" engaged in a contest to explain variation in the outcome. When aspects are examined as independent variables, which is the norm in conventional multivariate analysis, the impact of limited diversity is masked and obscured, as it would be in the present exam- ple. Variable-oriented researchers are typically unaware that their con- clusions incorporate assumptions about cases for which they have no evidence (e.g., minority members who are tenured). This issue per- vades empirical social science because limited diversity is the rule, not the exception, in the study of naturally occurring social phenomena.

I return to this topic—the impact of limited diversity on empirical generalizations—in subsequent chapters. As I show in part 2, the ex- amination of the impact of limited diversity is not limited to property

spaces constructed from categorical variables and conventional, crisp sets. These same concerns surface in the examination of aspects of cases that exhibit fine-grained variation by level or degree.

CONCLUSION

Seeing cases as configurations is a middle path between assuming that cases are "homogeneous enough" to equate their similarities, on the one hand, and attending to the specificity of each case, on the other. At the variable-oriented extreme, cases are studied primarily in terms of their component variables that, in turn, are treated as independent, analytically separable aspects. This transformation of cases disaggre- gates wholes and masks limited diversity. Referring back to table 3.3, it is clear that if the researcher were to study these variables as analytically separable and ignore their different combinations, then the department would appear to be much more diverse than it actually is. Further, as I have shown with a simple example, when researchers ignore limited diversity, assumptions about combinations of conditions that have not been examined can easily become embedded in empirical generaliza- tions.

At the case-oriented extreme, by contrast, cases are selected and studied because of their special significance, so much so that research- ers may lose sight of the larger context. For example, a researcher might conduct an in-depth study of the female untenured faculty in the de- partment with the goal of documenting their efforts to thrive in a large and diverse academic department. However, this researcher might fail to consider the impact of several basic features of this department: (1) all women, whether tenured or not, receive below-average pay, and (2) untenured faculty who are both minority and female work in a department that lacks tenured minority faculty members. In short, the general character of the department might be obscured in the attempt to document the experiences of a relatively narrow category of faculty.

Viewing cases as configurations is a not a panacea. There is no substi- tute for insight or for in-depth knowledge of research subjects. Still, by viewing cases as configurations it is possible to avoid many of the homogenizing assumptions of the variable-oriented approach without adopting the laser beam–like focus of the case-oriented approach. As I have shown, it is possible to study cases as configurations by extending Lazarsfeld and Barton's property space approach, which in turn, pro- vides a strong analytic foundation for constructing types and analyzing diversity.

FOUR

CAUSAL COMPLEXITY

CAUSAL COMPLEXITY AND THE STUDY OF DIVERSITY

Social diversity is manifested in many different ways. One of its clearest manifestations is in the variety of ways things happen. Choose almost any outcome that a social scientist might study and consider all the different ways it could come about. As an example, take democracy: In some countries, democracy developed through local institutions and matured gradually. In others, democratic institutions were borrowed wholesale from other countries, and political leaders established democracy more quickly. In still others, democracy was imposed from the outside all at once, sometimes by a colonial power, sometimes by the victor in a war. Causal heterogeneity of this type is not limited to large-scale, historically emergent outcomes like democracy. At a more mundane level, consider how many different ways there are for an individual to get ahead (or to fall behind) or how many different ways there are to strike up a conversation (or to avoid interacting with others). It follows that one important way to address social diversity is to pay close attention to the variety of ways a common outcome is reached—that is, to attend to causal complexity.

Of course, social scientists like to generalize about causes. If they can, they try to identify powerful, generic causes that are relevant to broad populations. At the individual level, the concept of self-interest is one such cause; so is the idea that it is important for individuals to accumulate human capital. At the macrosocial level, economic development is a popular general cause. Social scientists often argue that economic development fosters democracy, political stability, civic culture, tolerance, and so on. The urge among social scientists to find general causes and identify broad patterns is pervasive, and it is usually healthy.

After all, the opposite of focusing on broad patterns and general causes is to see each instance as irreducibly unique—different in almost every way from all other instances. But there can be too much of a good thing. Social scientists' preference for broad generalizations often leads them down the path of excessive abstraction and away from understanding diversity. For example, the finding that democratic stability is very often linked to economic development is important. However, our understanding of democracy is greatly deepened when we uncover the different ways democratic stability can be forged, even in poor countries (Rueschemeyer, Stephens, and Stephens 1992), and the different ways it may break down, even in advanced countries (Berg-Schlosser and De Meur 1994).

In this chapter I explore the link between social diversity and causal complexity. Along the way, I address various aspects of causation and the problem of establishing causal generalizations that strike a balance between complexity and generality. My goal is not to criticize conventional views of causation in the social sciences. Some scholars, for example, reject the use of causal imagery to describe cross-case patterns (e.g., the argument that "more education leads to higher income" based in part on a cross-case correlation between these two variables). Rather, I am concerned with assumptions about how causes operate—for example, the idea that causes are additive in their effects—that typically undergird quantitative analyses of cross-case patterns.

My starting point is the case study and the problem of generalizing from a single case. I proceed from a brief examination of the case study to basic ideas about causation, especially the concepts of causal necessity and sufficiency. I address the necessity and sufficiency of single causes and then augment the framework to address combinations of causes, multiple combinations of causes, and finally the issue of probability and inference in the study of necessity and sufficiency. I conclude by contrasting the understanding of causation advocated here—emphasizing causal complexity as an important key to understanding social diversity—with the understanding of causation that is implicit in conventional quantitative analysis.

My main argument is that researchers interested in diversity, especially as it is manifested in causal complexity, should avoid, as much as possible, making simplifying assumptions about the nature of causation. Specifically, I argue that researchers should avoid *assuming* that the individual causes they examine are either necessary or sufficient for the outcomes they study. I focus on techniques appropriate for studying causes that are sufficient only in combinations.

NECESSITY, SUFFICIENCY, AND CAUSAL GENERALIZATION

While the case study almost by definition offers little basis for causal generalization, it has the advantage of providing the investigator intensive knowledge of a case and its history and thus a more in-depth view of causation. Case-study researchers are able to triangulate different kinds of evidence from a variety of different sources in their attempts to construct full and compelling representations of causation in the cases they study. In short, case studies maximize validity in the investigation of causal processes.

Maximizing validity carries costs, however. While the case study is an excellent research strategy for studying "how" something comes about, it does not provide a good basis for assessing the generality or the nature of the causation the researcher identifies. For example, there is no way to tell from a single case if the causes identified by the researcher are either necessary or sufficient for the outcome in question. In the typical case study all causes identified as important by the researcher appear to be necessary conditions, for they are all present in the case in question and are typically understood as combinatorially decisive for the outcome under investigation. But are they truly necessary conditions? This question can be answered only by examining other, comparable instances of the outcome. There are other important issues that cannot be addressed in a case study. Examples of these issues include the size of the set of cases that is relevant to the causal pattern or process observed in the case study, the relevant boundary or scope conditions for the causal process observed by the case-study researcher, and whether the researcher identified all relevant causal factors.

A researcher studying a specific labor strike, for example, might conclude that the strike occurred as a result of worker opposition to the introduction of new machinery designed to increase worker productivity. In the report of the case study of this strike, the researcher might present a very compelling array of evidence supporting this conclusion, based on observations, interviews, experience working on the production line, the study of relevant documents, an analysis of the history of the plant in question, references to the existing literature on strikes, and so on. However, even with this compelling evidence in hand, it is impossible to know if it is reasonable to extend the researcher's argument to other cases. Claims about *generalizability*—that is, the portability of findings to other, comparable cases and settings—in social science usually rest on some form of *cross-case* analysis or evidence. That is, such statements usually are based on the examination of many cases, judged to be similar enough to warrant pooling the evidence they provide.

There are several different ways to address the generalizability of empirical findings.[1] I address two aspects of generalizability here because these two are of special importance to the study of social diversity as reflected in causal complexity. The first involves the *necessity* of the cause identified by the researcher. In the example just presented, one could ask: Are *all* strikes preceded by the introduction of new technology that is opposed by workers? Or, are strikes also prompted by other causes? The second aspect of generalizability of interest here involves the *sufficiency* of the cause in question. Is the cause identified by the researcher by *itself* capable of producing the outcome? In other words, whenever this cause occurs in like settings, regardless of what other causally relevant factors are present, will there be a strike? Or are there other essential ingredients—which may have escaped the researcher's grasp—that must be combined with the cause in question in order for a strike to occur?

To assess necessity, the researcher must work backward from instances of the outcome to the identification of relevant causes. As noted, a necessary cause must be present for the outcome in question to occur. Thus, all instances of the outcome should be preceded by the cause or exhibit the cause in some way. If another researcher were to identify strikes not preceded by the introduction of new technology, then this would challenge the claim that the cause identified by the first researcher is a necessary cause. Assessing necessity thus involves searching for instances of the outcome (in this example, strikes), and then assessing whether all instances of the outcome agree in displaying the same cause. If it can be shown that, in fact, all instances of the outcome share the same antecedent condition, then the researcher has established that the antecedent condition may be necessary for the outcome.[2]

1. Here I depart from conventional approaches, where generalizability would be conceptualized in terms of probabilities. For example, in conventional approaches a researcher might ask, What impact does the introduction of new technology opposed by the workers have on the *probability* of a strike? Does it increase the probability, and if so, by how much? It is not conventional to address causal necessity or sufficiency within a probabilistic framework. I return to this issue after establishing important analytic consequences of considering causal necessity and sufficiency in social research.

2. In the course of making this assessment, the researcher may discover that there are preconditions that were hidden from view in the first case. For example, the researcher might "discover" that his or her argument about strikes really only applies in industries that have been unionized for a long period of time.

Of course, the researcher addresses only those causes that are theoretically or substantively relevant in some way. Most researchers know, for example, that it is trivial to point out that "air for workers to breathe" is a necessary condition for strikes.

Necessary causes are not always sufficient causes. Even if it can be shown that all known and relevant strikes were preceded by the cause identified by the case-study researcher, it would be erroneous to conclude on the basis of such evidence that whenever this cause occurs, a strike results. It may be the case, for example, that another, unidentified cause also must be present and that there are actually two (or more) causes that, when combined, provide sufficient conditions for a strike.

To assess the sufficiency of a cause, the researcher must determine whether the cause in question always produces the outcome in question. Evidence that there are instances of the cause not followed by the outcome challenges the researcher's claim that the cause in question is sufficient. The assessment of sufficiency, therefore, involves searching for cases that are similar to the present case with respect to the cause in question (in this example, the introduction of technology opposed by workers) and then assessing whether they agree in displaying the same outcome (strikes). If all instances of the cause result in strikes, then the researcher may argue that the cause is sufficient for the occurrence of a strike. Combined with the evidence in support of necessity described previously, the researcher could claim that the cause identified in the case study is a necessary and sufficient condition for strikes.[3]

The sequence of analytic steps just presented roughly reproduces the "indirect method of difference" outlined by John Stuart Mill ([1843] 1967) in *A System of Logic*. The indirect method of difference is a double application of his "method of agreement." First, the researcher searches for and examines instances of the effect to see if they all agree in displaying the same antecedent condition (i.e., the researcher assesses the *necessity* of a cause). Second, the researcher searches for and examines instances of the cause to see if they all agree in displaying the effect

(i.e., the researcher assesses whether the necessary cause identified in the first step is also *sufficient*). While the indirect method of difference does not provide the causal clarity of a laboratory experiment, it is an important way to assess causation when working with naturally occurring data—the usual situation for social scientists.

While Mill's indirect method of difference offers a relatively simple and straightforward research design, it is easy to be fooled about the complexity of causation by the sequential nature of the design just described. The design leaves the impression that the first task is to certify that a cause is necessary, while the second is to show that not only is the cause necessary but that it is also sufficient. This research plan thus assumes causal uniformity at the outset of the investigation because the search for common causes is predicated on the idea that instances of the outcome share basic similarities in how they came about. However, it is often the case that there are no necessary causes for an outcome—beyond those that simply establish the boundaries of the investigation (e.g., the presence of workers might be seen as a necessary but trivial condition for the occurrence of a labor strike).

In fact, causation is often complex in character because social phenomena are remarkable in their diversity, even phenomena that merit the same label, like *strikes*. Causes may be sufficient but not necessary, and they may be necessary but not sufficient. Furthermore, in many arenas of social scientific research causes may be neither necessary nor sufficient. In fact, this type of causation may be the most common form of social causation. While it is tempting to reject causes that are neither necessary nor sufficient as "not general" and therefore of only marginal interest to social science, to do so would be to deny the complexity and diversity of social phenomena, especially with respect to causation. Research strategies appropriate for uncovering and assessing social diversity must permit maximum causal complexity, and the most complex form of causation involves conditions that are neither necessary nor sufficient.

For elaboration of these ideas, examine table 4.1. It shows four types of causes. In cell 1 are causes that are both necessary and sufficient.

Table 4.1
Types of Causes

Cause	Sufficient	Not Sufficient
Necessary	cell 1	cell 2
Not necessary	cell 3	cell 4

3. These two very different assessments are often conflated or confused in social research. Qualitative researchers often proceed from a set of cases defined by the outcome, attempt to identify causes that are uniform across all such cases, and then present these uniform causes as necessary and sufficient conditions. As is evident from the preceding discussion, however, researchers who use this strategy succeed only in establishing necessary conditions for an outcome, and in most instances they establish only a subset of the necessary conditions. It is also clear that such work rests on an assumption of causal homogeneity—the expectation that the same causes operate in the same way in all instances of an outcome.

Causes of this type have the greatest empirical scope (because they apply to all relevant instances) and the greatest empirical power (because the cause by itself produces the outcome). However, such causes are very rare in the study of social phenomena. In cell 2 are causes that are necessary but not sufficient. These causes also have great empirical scope, but they lack the empirical power of cell 1 causes because they work only in conjunction with other causes. Sometimes, such conditions merely establish the scope of an investigation without specifying central causal mechanisms. Causes in cell 3 are powerful because they can act alone to bring about an outcome, but their empirical scope is limited because there are other causes that also produce the same outcome. Finally, causes in cell 4 are limited both in empirical scope and power, because they do not produce the outcome on their own nor are they always present as antecedent conditions.

It is possible to use simple logical statements to express the causal concepts elaborated in table 4.1. Suppose the investigator considers four possible causes of strikes: (1) the introduction of new technology opposed by workers ("technology"), (2) stagnant wages in times of high inflation ("wages"), (3) reduction in overtime hours ("overtime"), and (4) worker resistance to outsourcing portions of an existing production process ("sourcing"). Suppose also that a researcher has evidence on all relevant production sites. Consider the following logical statements expressing different possible findings from these cases, with respect to the impact of new technology:

$$\text{technology} \rightarrow \text{strikes}, \tag{1}$$

$$\text{technology} \cdot \text{wages} \rightarrow \text{strikes}, \tag{2}$$

$$\text{technology} + \text{wages} \rightarrow \text{strikes}, \tag{3}$$

$$\text{technology} \cdot \text{wages} + \text{overtime} \cdot \text{sourcing} \rightarrow \text{strikes}, \tag{4}$$

where midlevel dots indicates logical *and* and plus signs indicates logical *or*. The first equation indicates that the researcher found that all strikes were preceded by a single cause (the introduction of new technology) and that this single cause invariably provoked strikes. In this equation new technology is a necessary and sufficient condition for strikes. The second equation indicates that the researcher found that two causes combined to produce strikes: new technology and stagnant wages. In this equation technology is a necessary but not sufficient condition for strikes (as is stagnant wages). The third equation indicates that the researcher found that either of two causes, new technology or stagnant wages, cause strikes. In some instances strikes occurred because new technology was introduced; in others, strikes occurred because wages were stagnant. In this equation new technology is a sufficient but not necessary condition for strikes (again, as is stagnant wages). Finally, the fourth equation shows that the researcher found that two different combinations of conditions caused strikes: workers struck when new technology was introduced while wages were stagnant, and they struck when overtime hours were reduced in concert with an attempt to outsource portions of the production process. In this equation, new technology is neither necessary nor sufficient for strikes.

The fourth type of causation (shown in equation [4]) is the most complex of the four presented in table 4.1. In this equation, none of the four single causes (new technology, stagnant wages, reduced overtime hours, or outsourcing) is either necessary or sufficient. Consequently, none of the single causes acts as a cause in all instances of the outcome (i.e., is a general cause), and none is capable of producing the outcome on its own (i.e., is singularly sufficient).

From the perspective of variable-oriented research the ideal cause should be relevant to as many cases as possible (i.e., it should be general) and also capable of acting independently to effect or influence the outcome. When causes operate the same way in all cases, the investigator can say that the cases included in a study are uniform with respect to the cause in question—that they display "causal homogeneity." When a cause is capable of acting by itself on an outcome, the investigator can say that its effect is independent of the presence or absence (or degree of presence) of other conditions. When there are several different causes and each one has an independent effect on the outcome (i.e., the strength of its effect is not altered by the presence or absence or level of other causes), then the effect of each cause can be described as "additive."

The single causes represented in equation (4) possess none of these attributes. Yet, the type of causation indicated in equation (4) may be the most common form of social causation. Even if it is not the most common form, a social science that attends to social diversity should assume at the outset of any empirical study that the causal processes under investigation may in fact be this complex—they may involve causes that are neither necessary nor sufficient. Of course, empirical investigation may reveal otherwise. The research may show that single causes under investigation are necessary or sufficient. But these simpler forms of causation should not be assumed at the outset, as is commonly done in conventional quantitative social science (especially via the assumption of additivity). Necessity and/or sufficiency must be established through empirical analysis.

ANALYTIC STRATEGIES RELEVANT TO DIFFERENT TYPES OF CAUSES

Social scientists have been slow to recognize that different analytic strategies are relevant to the assessment of different kinds of causes. At the most basic level, it is important to recognize that the study of necessity works backward from instances of an outcome and is a search for common antecedent conditions. The study of sufficiency, by contrast, works forward from instances of a causal condition (or, as I show subsequently, instances of a combination of causal conditions) to see if these instances agree in displaying the outcome. I elaborate on these different emphases and their implications in the discussion that follows, using dichotomous, presence/absence conditions and a simple two-by-two table.

In conventional variable-oriented research, all four cells of the cross-tabulation of the presence/absence of an effect against the presence/absence of a cause are considered relevant to the investigator's argument. Basically, cases in the cells where the cause and the effect are present or where the cause and the effect are absent count in favor of the inference of a causal relationship, while cases in the two other cells count against it. This simple principle is the foundation of almost all quantitative analysis in the social sciences today, including Pearson's correlation coefficient, the computational foundation of conventional multiple regression analysis. As I show subsequently, however, the reasoning behind these calculations conflates the analysis of necessity and the analysis of sufficiency. Some errors of prediction violate sufficiency; others violate necessity. Furthermore, cases where both the cause and the effect are absent are not directly relevant to the assessment of either necessity or sufficiency. Most measures of association count cases in this cell as evidence in favor of a causal argument: the greater the number, the better. From the perspective of necessity and sufficiency, however, this common practice is misleading.

For illustration, first consider the pattern of empirical results that should accompany the finding that a cause is both necessary and sufficient (as detailed in cell 1 of table 4.1 and equation [1]). Recall that to assess necessity, the researcher must locate all relevant instances of an outcome and then assess whether they agree in displaying one or more antecedent conditions. Essentially, this assessment involves an examination of only those cases where the outcome is present. The key question is, Are there any cases where the outcome is present but the cause is absent? If there are such cases, then the test of necessity fails. The assessment of sufficiency, by contrast, involves only cases where the cause is present. The researcher examines instances of the cause and hopes to show that in all instances of the cause, the outcome is present. If there are instances of the cause without the outcome, then the test of sufficiency fails.

Table 4.2 shows a simple cross-tabulation of the presence/absence of a cause against the presence/absence of an outcome, consistent with a finding of *both* necessity and sufficiency. Essentially, the test of necessity involves only the first row of the table (cells 1 and 2). The test of sufficiency involves only the second column (cells 2 and 4). Note that cell 3 is not directly relevant to the assessment of either sufficiency or necessity; therefore, it does not matter how many cases are in this cell. More than likely, there would be many cases in this cell in the hypothetical study of strikes just described because the investigator has information on "all relevant production sites" and there would certainly be many that lacked both the cause and the outcome. Still, tests of sufficiency and necessity are not directly affected by cases in cell 3.

Next consider the situation where the researcher is interested only in assessing whether the cause in question is necessary (as detailed in cell 2 of table 4.1 and equation [2]). This situation often arises when the boundaries of the set of negative cases (i.e., cases not displaying the outcome) is unclear, arbitrary, difficult to pin down, or potentially infinite (e.g., the set of nonrevolutions in a study of revolutions). Thus, the investigation is restricted to positive cases and thus to the assessment of necessity. The relevant empirical pattern is shown in table 4.3. In this analysis, only the first row of the cross-tabulation is relevant. The researcher's objective is to show that there are no instances of the outcome lacking the cause (i.e., no cases in cell 1).

Next, consider the situation where the researcher is interested only in assessing whether a cause is sufficient (as detailed in table 4.1, cell

Table 4.2
Cause Is Necessary and Sufficient

Outcome	Cause Absent	Cause Present
Present	[cell 1] no cases	[cell 2] cases
Absent	[cell 3] not directly relevant	[cell 4] no cases

Table 4.3
Cause Is Necessary but Not Sufficient

Outcome	Cause Absent	Cause Present
Present	[cell 1] no cases	[cell 2] cases
Absent	[cell 3] not relevant	[cell 4] not relevant

Table 4.4
Cause Is Sufficient but Not Necessary

Outcome	Cause Absent	Cause Present
Present	[cell 1] not relevant	[cell 2] cases
Absent	[cell 3] not relevant	[cell 4] no cases

3 and equation [3]) and has no particular concern for necessity. This situation might arise in research where investigators believe that there are no necessary causes (of a nontrivial nature) for the outcome in question and that there is more than one way to generate the outcome.[4] The relevant empirical pattern is shown in table 4.4. In this analysis, only the second column of the cross-tabulation is examined. The key concern is to show that there are no cases in cell 4 (cause present, outcome absent). The researcher can effectively ignore information on cases where the cause is absent because they are irrelevant to the assessment of sufficiency. After all, the researcher is convinced that there are other causes that produce the outcome, so there probably should be cases in cell 1. Again, the number of cases in cell 3 is not directly relevant.

Finally, consider the situation in which the researcher believes that the cause in question is neither necessary nor sufficient. The analyst can anticipate finding cases in all four cells:

- where the cause in question is present and the outcome is absent (cell 4—because the cause by itself is not sufficient);
- where the outcome in question is present but the cause is absent (cell 1—because the cause is not necessary);
- where both the cause and outcome are present (cell 2—because there will be some instances of the single cause in which it is part of a sufficient combination of causes; see equation [4]); and
- where both the cause and the outcome are absent (cell 3).

4. "Trivial" necessary causes sometimes define the set of relevant cases (the population of observations).

In short, if the researcher believes that a cause is neither necessary nor sufficient but instead is part of one of several sufficient combinations of conditions (as in equation [4]), then the assessment of the relationship between a *single* cause and the outcome in question is of relatively little analytic value.

If, indeed, the most common form of social causation involves causes that are neither necessary nor sufficient, then the assessment of the cross-tabulation of a single cause with the outcome in question provides little useful information. Instead, as I show in the next section, the investigator should cross-tabulate the outcome against *combinations* of causes. This shift in analytic strategy is the first step on the road to the analysis of causal complexity, defined here as a situation where no single cause is either necessary or sufficient.

NECESSITY AND SUFFICIENCY APPLIED TO CAUSAL COMBINATIONS

Social phenomena rarely result from single causes. Think about the outcomes that usually interest social scientists. They study social phenomena such as racial and ethnic conflicts, marriages, strikes, elections, revolutions, social movements, divorces, corporate mergers, street gangs, and so on—the many things that people do together. In virtually every phenomenon just mentioned, the causation involved is typically conjunctural—the relevant causes must be combined in the same time and place to produce the social phenomenon or outcome in question. There is no "one cause" of racial conflict or "one cause" of divorce. Social phenomena typically result from a combination of conditions, and very often the same outcome will result from several different combinations.

Consider again the case-study researcher studying a specific labor strike. When case-study researchers focus on the causes that account for an outcome, they almost invariably cite a combination of causes, an intersection or conjuncture of causal forces in time and space. For example, instead of concluding that the strike resulted from a single cause (the introduction of new technology opposed by workers), more than likely the conclusion would be that several important causal conditions (low unemployment, a strong union, and poor communication between workers and management) accompanied the introduction of new technology, and it was this combination of conditions, this causal conjuncture, that provoked the strike. In essence, the researcher might argue that four causes (three describing important features of the setting and the fourth naming a precipitating event) combined to produce a qualitative outcome, the labor strike.

Just as it is possible to assess the necessity and sufficiency of a single cause, it is also possible to assess the necessity and sufficiency of a combination of causes. The procedures described in the previous section can be duplicated, except that the focus is on combinations of causes rather than on single causes. Again, imagine that the researcher has information on "all relevant production sites"—without worrying, for now, about how this set of cases might be constructed.

When a causal combination is *necessary* for an outcome, all instances of the outcome should exhibit the same combination of causal conditions. By examining instances of the outcome, it is possible to see if they share a specific combination of conditions (e.g., the four conditions identified in the hypothetical case study of the labor strike described in the preceding paragraphs). If they share a specific combination of conditions, then the combination can be treated as a necessary combination. Of course, it should be noted that if a combination of conditions is necessary for an outcome, then each single condition within the combination is also necessary for the outcome. This result follows logically because instances of an outcome that agree in displaying a combination of conditions also will agree in displaying each single condition in the combination.

When a causal combination is *sufficient* for an outcome, all instances of the causal combination should be followed by the outcome in question. By examining all relevant instances of the causal combination, it is possible to see if they agree in producing the outcome. If there are instances of the causal combination not followed by the outcome, then the test of the sufficiency of the causal combination fails.

For illustration, consider table 4.5. Once again, the assessment of necessity essentially involves the first row of the table (cells 1 and 2), while the assessment of sufficiency involves the second column of the table (cells 2 and 4). In the assessment of necessity, cell 2 is expected to contain all the positive cases; cell 1 should be void of cases. Cells 3 and 4 are not directly relevant to the assessment of necessity. In the assessment of sufficiency, again cell 2 is expected to hold relevant positive cases; cell 4 should be void of cases. Cells 1 and 3 are not directly relevant to the assessment of sufficiency. While there are likely to be cases in cell 3, the number of cases in this cell is not directly relevant to the assessment of either the necessity or the sufficiency of a causal combination.

Table 4.5
Assessing a Causal Combination

Outcome	Causal Combination Absent	Causal Combination Present
Present	[cell 1] key cell for assessing necessity	[cell 2] cases
Absent	[cell 3] not directly relevant	[cell 4] key cell for assessing sufficiency

Imagine the application of these procedures to "all relevant production sites" using the results of the more elaborate case study, described previously, for guidance. Recall that in this hypothetical case study the investigator identified four causes that combine to provoke a strike (the introduction of new technology opposed by workers, low unemployment, a strong union, and poor communication between workers and management). Suppose the analysis of the cross-case evidence on many production sites revealed that all cells have cases except cell 4. The researcher could argue, based on this evidence, that the causal combination in question is sufficient but not necessary for strikes. By contrast, the finding that all cells have cases except cell 1 would support the argument that the causal combination in question is necessary but not sufficient for strikes. If both cells 1 and 4 prove to be void of cases, the investigator could argue that the causal combination is necessary and sufficient for strikes.

As noted previously, if a specific combination of causes is necessary for an outcome, then each single cause within the combination is also necessary for the outcome. It follows that if researchers were to test the necessity of the component conditions one at a time, they would find that each single condition passes the test of necessity (i.e., a showing of no "cell 1" cases). Thus, one simple way to identify a necessary combination of conditions is to test the necessity of each causal condition one at a time and then combine all conditions that pass this test into a single expression. A necessary combination of conditions is nothing more than the intersection of all the individually necessary conditions. This same procedure does not work, however, for combinations of conditions that are sufficient but not necessary. The component conditions of a sufficient combination of conditions generally fail the sufficiency test when examined one at a time. As I explain in chapter 5, it is important to examine all possible combinations of relevant conditions when assessing sufficiency.

MULTIPLE COMBINATIONS OF CAUSES

One situation not yet addressed is the possibility that *different* combinations of conditions may produce the same outcome. Suppose that there are two combinations of conditions that produce strikes. The first is the one just described (low unemployment, a strong union, poor communication with management, and the introduction of new technology opposed by workers). The second involves a different combination of four causes: a weak union, low wages, a concerted effort by the union to boost membership, and a speedup in the pace of production following a sharp increase in the demand for the product manufactured by the workers. (This second combination of causes does not involve the introduction of new machinery or production technology, only a speedup of production using existing techniques.)

Because there are two combinations, the analysis of cross-case evidence would show that neither combination is necessary. In both analyses there would be cases in cell 1 because there is plural causation and thus instances of the outcome caused by the other combination. In fact, because the two causal combinations display no common causal conditions, no single cause would pass the test of necessity. However, both combinations would pass the test of sufficiency. In both cross-case analyses, cell 4 would be void of cases, while cell 2 would contain the instances explained by the causal combination in question. This pattern of results is shown in table 4.6.

Table 4.6 demonstrates that the assessment of sufficiency is not undermined by the existence of multiple combinations of causes. Because the researcher focuses only on the results for the combination of causes in question (i.e., only cells 2 and 4) when assessing sufficiency, the impact of other causal combinations is excluded from the assessment. Thus, as table 4.6 shows, it is possible to assess the sufficiency of combinations of causal conditions one at a time—in isolation from one another.

This conclusion is important because of its implications for the study

Table 4.6
Assessing Multiple Causal Combinations

Outcome	First Causal Combination Absent	First Causal Combination Present
Present	[cell 1] cases explained by other causal combination	[cell 2] cases explained by first causal combination
Absent	[cell 3] not directly relevant	[cell 4] no cases

of social diversity as manifested in causal complexity. If, as I have argued, we live in a world of great causal complexity, then a common pattern will be for outcomes to result from different combinations of causal conditions. When there are different combinations of causes with no overlap in the single causes included in each combination (as in the present example), then no single cause is either necessary or sufficient. While it might seem that causation this complex should befuddle analytic social science, it is clear from the example just presented that the analysis of the sufficiency of causal combinations can proceed in a straightforward manner, with the key question being the distribution of cases in the second column of the cross-tabulation.

This analytic strategy—the examination of the sufficiency of combinations of causal conditions—is the preferred tactic for the study of causal complexity. This strategy makes no assumptions about the empirical scope or power of the causes examined in social research. It assumes neither causal uniformity (the same causes are involved in each instance of the outcome) nor causal additivity (causes act independently on the outcome) in the analysis of cross-case empirical evidence. The analysis of relevant empirical evidence may reveal these patterns (i.e., causal homogeneity or additivity), but they are not assumed to exist at the outset of the investigation, as in most quantitative social research today.

Finally, suppose a researcher assessed the two causal combinations just described and found them both sufficient. It would be reasonable as a next step to assess whether these two sufficient causal combinations account for all instances of strikes found in this particular set of cases. The relevant results are shown in table 4.7. If all instances of the outcome are covered by the two causal combinations, then there will be no cases in cell 1. In this table the two combinations are understood as alternate causal conjunctures or paths. The causal conditions specified in the last column of table 4.7 (i.e., "first or second causal combination present") can be interpreted as a logical statement, and the satisfac-

Table 4.7
Joining Multiple Causal Combinations

Outcome	Neither Causal Combination Present	First *or* Second Causal Combination Present
Present	[cell 1] no cases	[cell 2] cases explained by the causal combinations
Absent	[cell 3] not directly relevant	[cell 4] no cases

tion of this formula can be viewed as a necessary and sufficient condition for the outcome. The statement is

technology·employment·strong·communication
+ speedup·wages·weak·membership → strikes,

where "technology" equals introduction of new technology opposed by workers, "employment" equals low unemployment, "strong" equals strong union, "communication" equals poor communication between workers and management, "speedup" equals production speedup, "wages" equals low wages, "weak" equals weak union, "membership" equals union membership drive, midlevel dots equal logical *and*, and plus sign equals logical *or*. Essentially, this equation describes a pattern of "multiple conjunctural causation" (Ragin 1987). Different combinations of conditions are linked to the same outcome. As presented, the two causal combinations listed in the formula are mutually exclusive because the first specifies the presence of a strong union as a condition, while the second specifies the presence of a weak union as a condition.

THE IMPACT OF LIMITED DIVERSITY ON CAUSAL COMPLEXITY

The main focus of this chapter is on causal complexity, especially the possibility that no single cause may be either necessary or sufficient for an outcome. The main recommendation for dealing with causal complexity is to focus on the sufficiency of combinations of causal conditions. However, it is important to recognize that there are some situations in which the existence of necessary conditions will be obscured if scholars focus exclusively on the analysis of sufficiency. Generally, when the diversity of combinations of causal conditions is limited empirically (see chapter 3), researchers may overlook necessary conditions. Thus, it is always important to examine the impact of limited diversity on any conclusion that is drawn from analyses of necessity and sufficiency. Researchers must beware that their conclusions may incorporate assumptions about hypothetical combinations of causal conditions (see also chapter 3).

For a very simple illustration of the general problem, consider a researcher interested in the emergence of "generous" welfare systems in advanced industrial democracies. Suppose this researcher examines only two causal conditions: the presence/absence of strong "left parties" (e.g., social democratic parties in Scandinavia) and the presence/ab-

Table 4.8
Limited Diversity and the Analysis of Necessity and Sufficiency

Row No.	Strong Left Party	Strong Unions	Generous Welfare State	Number of Countries
1	Yes	Yes	Yes	6
2	Yes	No	No	7
3	No	No	No	5
4	No	Yes	?	0

sence of strong centralized unions. Suppose further that the empirical evidence follows the pattern presented in table 4.8. Using these two causal conditions as the main dimensions of the relevant property space, the table shows that only three of the four cells in this property space have cases. Thus, the diversity of cases, with respect to the relevant causal conditions, is clearly limited. There are no countries, in this hypothetical study, that combine strong unions and weak left parties.

Viewed from the perspective of causal sufficiency (where researchers work forward from causal conditions to outcomes), two causal arguments pass the sufficiency test for generous welfare systems: strong unions and the combination of strong unions and strong left parties. Logic and parsimony dictate that having strong unions, by itself, is sufficient for generous welfare systems; the combination of strong unions and strong left parties thus can be dropped (see chapter 5 for further explication of this elimination). Thus, viewing the evidence exclusively in terms of sufficiency leads to the conclusion that having strong unions is sufficient by itself for having a generous welfare system. Furthermore, the cross-tabulation of the presence/absence of this condition against the presence/absence of the outcome reveals a perfect relationship, demonstrating that it is also a necessary condition.

But what if necessity had been examined first? From the perspective of necessity, all instances of the outcome agree in displaying both strong unions and strong left parties. Thus, these two conditions are both necessary, using the procedures outlined previously for assessing necessity. Furthermore, when these two necessary conditions are tested for sufficiency, as a combination of conditions, it is apparent they pass this test as well. Viewed from the perspective of necessity first, therefore, the proper conclusion is that strong left parties and strong unions are jointly necessary and sufficient for generous welfare systems. Thus, it is clear from this simple example that the sufficiency test, described previously, obscures a necessary condition—having strong left parties.

Which conclusion is correct? In general, if a causal condition central to a researcher's argument can be shown to be a necessary condition (i.e., present in all instances of the outcome), then it should be evaluated as such. That is, once researchers have identified causal conditions that are uniform across instances of the outcome, then they should evaluate whether these conditions indeed make sense as necessary conditions, using theory, substantive knowledge, and whatever auxiliary evidence is at their disposal (e.g., in-depth knowledge of cases). If, in the end, this evaluation confirms the understanding of these causes as necessary, then they should be retained as necessary conditions. That is, such causes should not be subsequently eliminated simply because they appear to be logically redundant from the perspective of causal sufficiency. Thus, as a general rule, researchers should check for necessary conditions *before* they conduct sufficiency tests. Any causal condition that passes the test of necessity and that also passes the investigator's "plausibility" test should be retained as a necessary condition (see also chapter 5).

Observe that the two tests, necessity and sufficiency, give seemingly inconsistent answers in this example precisely because of the limited diversity of cases with respect to the combination of causes they display (see table 4.8). There are no cases that combine strong unions and weak left parties. If such cases existed and they all had generous welfare systems, then the researcher could more safely conclude that having strong unions, by itself, is necessary and sufficient for generous welfare systems. If, however, such cases existed and they lacked generous welfare systems, then the researcher could more safely conclude that strong unions must be combined with strong left parties for generous welfare systems to occur. Another way to say the same thing is to observe that, given limited diversity, no matter which conclusion the researcher presents, it involves statements (and thus assumptions) about combinations of conditions that have *not* been observed. The conclusion that strong unions and strong left parties must be combined for generous welfare systems to emerge assumes that if there were cases with strong unions and weak left parties, then these cases would not have generous welfare systems. Likewise, the conclusion that having strong unions, by itself, is enough assumes that if countries with strong unions and weak left parties existed, they would have generous welfare systems.

Conventional quantitative social science, for the most part, does not recognize the impact of limited diversity. In this approach, existing combinations of conditions are typically seen as the relevant universe of conditions (i.e., values are usually seen as "fixed"). Furthermore, these techniques usually privilege parsimony: the most parsimonious model is treated as the best. Thus, a quantitative analysis of the evidence in table 4.8 would lead quickly to the conclusion that having strong unions is the single cause of generous welfare systems. After all, the correlation between strong unions and generous welfare systems is perfect, while the correlation between strong left parties and generous welfare systems is weak. Why go any further? In fact, however, as I have just shown, the more parsimonious conclusion (i.e., where there are strong unions, generous welfare systems emerge) involves assumptions about the outcomes that hypothetical combinations would display, if in fact they did exist. Many implicit assumptions are hidden in the typical quantitative analysis, not only from investigators, but also from their audiences.

What is the most prudent course? In general, as I have argued, researchers should test for necessity before they examine sufficiency. That is, they should examine instances of the outcome to see if they agree on any of the causal conditions specified by the researcher. Of course, it is important for the researcher to further evaluate any conclusions from such assessments with auxiliary evidence. It is also important to recognize that if diversity is limited, any conclusion that investigators reach regarding causation involves assumptions about combinations of conditions for which researchers lack evidence.

Finally, it is worth emphasizing that limited diversity is the rule, not the exception, in the study of naturally occurring social phenomena. It poses analytic problems regardless of which technique investigators apply to their data.

USING PROBABILISTIC CRITERIA TO ASSESS NECESSITY AND SUFFICIENCY

For readers familiar with empirical social research, the discussion of necessity and sufficiency presented so far may seem fanciful. How often do social researchers, especially those working with large numbers of cases, find clear evidence of necessity or sufficiency in their cross-tabulations? All too often, cells 1 and 4 have plenty of cases in them. The idea that social scientists might find one or both of these cells empty seems preposterous, especially given how things usually turn out. Even when the analysis is limited to the examination of the sufficiency of combinations of conditions—the preferred tactic in the study of social diversity—the probability of finding cell 4 completely void of cases seems quite low.

There are many good reasons researchers should expect to find cases in cells 1 or 4—cases that challenge claims of necessity or sufficiency. The foremost reason is the fact that there is a lot of randomness in human affairs, as well as nonrandom factors that simply escape the purview of our theories. Consider again the example of strikes that occur in response to the introduction of technology opposed by work-ers. Suppose a charismatic worker makes a stirring speech on the behalf of the new technology and convinces workers not to go on strike, but instead to give the new production techniques a try. Suppose a flood closes the plant for a month, and then workers flock back to work, eager for overtime bonuses, when the waters recede. Suppose anar-chists have infiltrated the union leadership, and the rank-and-file mem-bers refuse to follow any of their recommendations, no matter how sensible they may seem. There are many such minor, obscure, or ran-dom factors that might interfere with the expected connection between a cause and an effect. It is virtually impossible to construct social scien-tific models that take account of every possible factor that might influ-ence some action or outcome. It comes as no great surprise, then, that many cases that "should" be in cell 2, at least from the perspective of the researcher, often end up in cells 1 or 4.

Consider also the fact that the data social scientists use are often imperfect and full of errors. Researchers studying strikes at "all relevant production sites" might make many mistakes in their observations, and the events themselves may be ambiguous. Also, as shown in chapter 2, it is all too easy to "misconstruct" a population. Consider as well the problem of deciding when new technology has been introduced. How new is "new"? What constitutes minor versus major changes in technol-ogy? Is the social scientist conducting this research qualified to judge production technologies? Consider the problem of coding strikes. What if a vote is taken in favor of a strike, but the action gradually crumbles on the very first day? Was it a strike? What if there is no strike vote, but a majority of workers call in sick when the new technology is intro-duced and production is shut down until employers reinstitute the old production techniques. Was it a strike?

Beyond these common problems of interpretation are simple human errors. What if the research is sloppy and coding forms are misread, misplaced, or misinterpreted? What if uncorrected data forms are acci-dentally mixed in with the corrected data forms? What if the assistant in charge of entering the data into the computer just had a fight with his or her lover? Cases can very easily find their way into the "wrong" cells. It is a wonder that social scientists ever produce "findings."

The procedures outlined in the previous sections for assessing neces-sity and sufficiency, therefore, must be modified to take these troubling aspects of social data—error, chance, randomness, and other factors—into account. In short, these common data and evidence problems pro-vide a very strong motivation to employ analytic techniques that make some use of probability theory, especially techniques that address the problem of drawing inferences from imperfect evidence.

To make the discussion of these techniques more manageable, I limit it to the problem of using probabilistic criteria to make inferences about the sufficiency of combinations of conditions. However, the procedures I present for incorporating probabilistic criteria into the assessment of sufficiency also can be incorporated into the assessment of necessity. The only difference is that when studying necessity the investigator selects cases with the same outcome and examines whether they display the same causal condition, instead of selecting cases with the same cause or combination of causes and evaluating whether they display the same outcome.

THE QUASI-SUFFICIENCY OF CAUSAL COMBINATIONS

Rather than impose absolute standards in all investigations (i.e., the rule that cell 4 of table 4.6 should be completely void of cases in the assessment of sufficiency), researchers also can make inferences about sufficiency using probabilistic methods.[5] For example, a researcher might be interested in whether the evidence supports the claim that a particular combination of causes is "usually sufficient." He or she would evaluate cases with the causal combination in question to see what pro-portion display the outcome. If this proportion is significantly greater than, say, .65, then the causal combination in question might be labeled "usually sufficient" for the outcome. Or, a researcher might be inter-ested in assessing whether the evidence supports the claim that a causal combination is "almost always" sufficient. If significantly greater than .80 of the cases displaying the combination of causes also manifest the outcome, then the researcher might claim that the causal combination is "almost always" sufficient. In short, it is possible to assess the *quasi-sufficiency* of causal combinations using linguistic qualifiers such as "more often than not" (.5), "usually" (.65), and "almost always" (.80)

5. Dion (1998) shows how to use Bayesian probability theory to assess neces-sity. The argument I present in this section extends his approach, applying it to the sufficiency of combinations of conditions, using conventional probability theory.

and applying formal statistical tests using these benchmark proportions.

Of course, the translation of linguistic qualifiers to proportions is not exact but approximate. The selection of the precise benchmark is to some extent arbitrary, but it must be made explicit by the researcher and thus be open to debate. The benchmark proportion can be varied, and results using different benchmarks can be compared. For example, the researcher might want to compare results using a .65 benchmark proportion ("usually sufficient") with results using a .50 benchmark ("sufficient more often than not") or a .80 benchmark ("almost always sufficient").[6]

Not only does the researcher specify a particular benchmark proportion of successes that the observed proportion of successes must exceed, he or she also must assess whether the observed success rate is *significantly* greater than the benchmark, using probabilistic criteria. That is, the researcher must conduct a formal test of the hypothesis that the observed proportion is greater than the benchmark proportion, using the null hypothesis that the observed proportion is either the same as (i.e., statistically indistinguishable) or less than the benchmark proportion. Essentially, this test involves estimating the standard error of the benchmark proportion (based on the number of cases involved in the assessment) and then using this standard error to evaluate the size of the positive gap separating the observed proportion from the benchmark proportion.

How does this assessment proceed? Again, consider the researcher studying strikes. The researcher believes that one of the sufficient combinations of conditions for the occurrence of strikes involves a combination of four factors: low unemployment, a strong union, poor communication between workers and management, and the introduction of new production technology opposed by workers. Suppose there are 91 instances of this combination of conditions and 70 of these instances result in strikes. The actual success rate is 70/91 or .769 (76.9%). Suppose the researcher wants to tests the hypothesis that this combination of conditions is "usually" sufficient for strikes, using .65 to define this benchmark. The observed proportion (.769) is greater than the

6. Bruce Western (1998, personal communication) points out that these benchmarks would not work for outcomes with very low baseline probabilities. To address such phenomena, much lower benchmarks must be used. However, the lower the benchmark, the greater the analytic distance to the concepts of necessity and sufficiency.

.65 used as the benchmark proportion, but is it significantly greater than .65?

To assess this difference, the researcher can use a simple z test for the difference between an observed proportion and a "population" proportion. The benchmark proportion provides the population proportion (.65); the observed proportion is .769. The general format for this test is described by Hays (1981:211–214):

$$z = \frac{(P - p) - \frac{1}{2N}}{\sqrt{\frac{pq}{N}}},$$

where P is the observed proportion, N is the number of cases displaying the causal combination, p is the benchmark proportion, and q equals $1 - p$. Essentially, this formula assesses the degree to which the observed proportion exceeds the benchmark proportion relative to the standard error of the benchmark proportion. The calculation is predicated on having an observed proportion that is larger than the benchmark proportion and includes a correction for continuity (see Hays 1981:214). Notice that according to the formula, the greater the gap between the observed proportion and the benchmark proportion, the larger the z value. Also, the greater the number of cases, the larger the z value because a large number of cases will result in a small denominator in the formula.

Using the hypothetical data just described, the calculation of the relevant z value,

$$z = \frac{(.769 - .65) - \frac{1}{2 * 91}}{\sqrt{\frac{.65 * .35}{91}}},$$

gives $z = 2.27$.

The next step is to evaluate the z value just calculated. How large must z be to assure the researcher that the observed proportion is significantly greater than the benchmark proportion? If the researcher is willing to take a one-in-twenty chance of reaching wrong conclusions about the null hypothesis (i.e., 95% confidence or .05 significance), then the z value must be greater than 1.65 (the z score corresponding to a one-tailed α of .05). If the researcher is willing to take a one in ten chance of being wrong, then the z value must be greater than 1.28 (the z score corresponding to a one-tailed α of .10). It is clear in this hypothetical analysis that the test of sufficiency succeeds, using either .05 or .10 significance. The gap between the observed proportion of successes (.769) and the benchmark proportion (.65) is great enough to provide clear evidence that the causal combination in question is "usually" sufficient for the outcome (strikes).

When the number of cases displaying the causal combination is 30 or fewer, a binomial probability test should be used instead of the z test. After all, the z test is a large-N approximation of the binomial test. Essentially, the binomial test assesses the probability of observing a specific range of "successful" outcomes, given an expected probability of success, which in turn is provided by the benchmark selected by the investigator (e.g., "sufficient more often than not" or .5). For example, a researcher might assess the probability of observing four or more successes in six trials, when the underlying probability of success is .5 (as in a coin toss). The binomial probability formula is

$$\binom{N}{r} p^r q^{N-r},$$

where N equals the number of cases displaying the causal combination, p equals the benchmark proportion, $q = 1 - p$, and r equals the number of cases displaying the outcome. Note that this formula is applied not only to the observed frequency of cases with the outcome but also to all frequencies that are superior to the observed frequency. These probabilities are then summed.

For purposes of illustration, assume that the researcher studying strikes observes only 13 instances of a causal combination, and 10 of these instances also display the outcome (strikes). In this analysis, the observed proportion of successes is the same as in the previous example ($10/13 = 70/91 = .769$), but now the number of cases is much smaller ($N = 13$). Essentially, the binomial test asks, What is the probability of observing 10 *or more* successes, out of 13 trials, if the underlying probability of success is the benchmark proportion? To maintain continuity with the earlier example, I use an underlying probability of .65 (i.e., "usually sufficient") in the demonstration of the binomial test. Using the binomial formula just described the probability of 10 successes is .1651; the probability of 11 successes is .0836, the probability of 12 successes is .0259; and the probability of 13 successes is .0037

(Hays 1981:647–51). In short, the probability of 10 or more successes in 13 trials is .2783.

This probability is far in excess of most conventional significance levels for this type of evidence (.05 and .10). It exceeds even the .20 level, which is used occasionally in small-N research. Thus, the researcher in this example would refrain from making any inference about sufficiency. That is, the researcher could not claim that the evidence supports the conclusion that the causal combination is "usually sufficient" for the outcome (using .65 as the appropriate benchmark proportion). This example shows that sufficiency tests are strongly influenced by the number of cases displaying the combination. With fewer cases, the researcher has less confidence that the observed proportion (.769) is superior to the benchmark proportion (.65).

To make the problem interesting, assume the researcher again finds only 13 instances of the causal combination in question and, again, only 10 display the outcome (strikes). Instead of using the benchmark proportion for "usually" sufficient (.65), the researcher decides to use a more generous benchmark, "sufficient more often than not" (.5) and sets the significance level at .05. Summing the relevant probabilities from the binomial test yields a probability of .0461, which indicates that the probability of observing 10 or more successes, when the underlying probability is .5, is smaller than the significance level, .05. Thus, the researcher could argue that even though the evidence does not lend support to the claim that the causal combination in question is "usually sufficient," it does lend support to the claim that the causal combination is "sufficient more often than not."

Finally, consider a common small-N situation: the pattern of results is "perfect" but the number of cases is very small. In this example, the researcher locates only four instances of the causal combination in question, but all four instances display the outcome in question. From a "veristic" as opposed to a probabilistic perspective, perfect sufficiency has been established: there are no "cell 4" cases. (A veristic evaluation of the evidence is a search for disconfirming cases. If any negative cases are found, even one, the test fails.) But how does it pan out using probabilistic criteria? Is an observed proportion of 1.0 based on four cases strong enough evidence to support the claim that the causal combination is "usually sufficient" ($p = .65$)? The binomial calculation yields a probability of .1785. These results show that even though there are no cell 4 cases (see table 4.6), we cannot reject the null hypothesis that the observed proportion (1.0) is the same as or less than the benchmark

Table 4.9
Number of Consistent Cases Needed to Pass Probabilistic Test of Sufficiency or Necessity for Different Ns

	Benchmark Proportion								
	.50			.65			.80		
N	α = .10	α = .05	α = .01	α = .10	α = .05	α = .01	α = .10	α = .05	α = .01
4	4	...	...	...	...	...	...	...	...
5	5	5	...	...	...	...	...	...	...
6	6	6	...	...	...	...	...	...	...
7	6	7	7	7	7	...	...	...	...
8	7	7	8	8	8	...	...	...	...
9	7	8	9	9	9	...	...	...	...
10	8	9	10	9	10	...	...	...	...
11	9	9	10	10	11	11	11	...	...
12	9	10	11	11	11	12	12	...	...
13	10	10	12	12	12	13	13	...	...
14	10	11	12	12	13	14	14	14	...
15	11	12	13	13	14	15	15	15	...
16	12	12	14	14	14	15	16	16	...
17	12	13	14	14	15	16	17	17	...
18	13	13	15	15	16	17	17	18	...
19	13	14	15	16	17	18	18	19	...
20	14	15	16	17	17	19	19	20	...
21	14	15	17	17	18	19	20	21	21
22	15	16	17	18	19	20	21	21	22
23	16	16	18	19	20	21	22	22	23
24	16	17	19	20	20	22	23	23	24
25	17	17	19	20	21	22	23	24	25
26	17	18	20	21	22	23	24	25	26
27	18	19	20	22	23	24	25	26	27
28	18	19	21	22	23	25	26	27	28
29	19	20	22	23	24	26	27	28	29
30	19	20	22	24	25	26	28	28	30

proportion (.65), using either .05 or .10 significance. Thus, when using probabilistic criteria, "perfect" evidence based on small *Ns* (four confirming cases, no disconfirming cases) may not be strong enough to support the argument that a causal combination is "usually sufficient." It should be noted, however, that the evidence just described (four positive cases, no negative cases) does support the claim that the causal combination is "sufficient more often than not" (benchmark = .5) using a significance level of .10 (binomial probability = .0625).

Table 4.9 presents the number of positive cases (i.e., the number of cases in cell 2 of table 4.6) needed to pass sufficiency for different size *Ns* (i.e., the number of cases in cells 2 and 4 of table 4.6 combined). For each *N* (shown in column [1]), the number of positive cases needed to pass nine different tests is shown, using three sufficiency benchmarks (.50, .65, and .85) and three significance levels (.10, .05, and .01). The table shows both the difficulty of passing probabilistic sufficiency tests when *Ns* are very small and the moderating effect of both lower benchmarks and more lenient significance levels. The calculations presented in table 4.9 are all based on the binomial test. (These figures apply to probabilistic tests of necessity as well; see chapter 8.)

The incorporation of probabilistic criteria sketched here offers an important alternative to the veritistic standard depicted in table 4.6—"no cell 4 cases allowed." Using the probabilistic techniques just described, it is possible to assess causal complexity while making allowances for the randomness and error inherent in both human social life and the work of social scientists. Also, as I have just demonstrated, by incorporating probabilistic criteria it is possible to take advantage of linguistic qualifiers such as "more often than not" and "usually" and thereby enlarge and sharpen the assessment of the sufficiency of causal combinations. Using the concept of quasi-sufficiency, it is possible as well to expand the discourse surrounding the nature of causation and establish an important new tool for the study of causal complexity and social diversity.

ASSESSING SUFFICIENCY: A SUMMARY OF THE DIFFERENT TESTS

The assessment of sufficiency can take either one of two general forms: veristic, with no tolerance for discordant outcomes among the cases conforming to a causal combination, or probabilistic, using benchmarks and linguistic qualifiers, as just demonstrated. In *The Comparative Method* and most applications of the techniques I present in that book, the assessment of the sufficiency of causal combinations is veristic: In order to be considered sufficient for an outcome, *all* the cases conforming to a particular causal combination must display the outcome in question. When the number of relevant cases is small, as in most macrosocial research, this method is the only one available. The other general type of sufficiency test is probabilistic. There are two versions of this test. The first is the small-*N* version based on the binomial probability test, used with *N*s of 30 or fewer cases. The second is the large-*N* version based on the *z* test for the significance of the positive difference between a sample proportion and a benchmark proportion, used with values of $N > 30$.

It should be noted that the veristic approach may be modified to include evaluation of the strength of the evidence (see, e.g., Ragin 1995). As already explained, if a causal combination includes *any* negative cases of the outcome, it fails the veristic test of sufficiency. In addition, the investigator may establish a *frequency threshold* for the number of positive instances. If a causal combination embraces one positive instance of the outcome and zero negative instances, does the evidence support the claim that the causal combination in question is sufficient for the outcome? Is two positive instances enough? How many does it take? The researcher must justify the frequency threshold used to evaluate sufficiency in each investigation. In some studies, especially small-*N* comparative studies of large-scale macrosocial processes and events, a claim of sufficiency may be based on a single positive instance. This standard is the de facto norm in comparative case-study research. In studies with moderate-sized *N*s, more positive instances may be required. Thus, there are two forms of the veristic test of sufficiency, one involving no frequency threshold (in essence, a threshold of one positive case), the other incorporating a frequency threshold specified by the investigator. Table 4.10 summarizes the four major types of sufficiency tests. (These four types apply to tests of necessity as well.)

Table 4.10
Types of Sufficiency Tests

Veristic: No Disconforming Cases Allowed		Probabilistic: Disconfirming Cases Permitted; Researcher Sets Benchmark and Significance Level	
No use of frequency criteria: One case is enough to pass test	Use of frequency criteria: Researcher sets threshold for pass	Small-*N*: Use of binomial test with 30 or fewer cases	Large-*N*: Use of *z* test with greater than 30 cases

CONTRASTS WITH CONVENTIONAL, VARIABLE-ORIENTED PRACTICES

Social scientists generally avoid discussing causal complexity. Consequently, they rarely have much to say about either necessity or sufficiency. The reasons they avoid these topics are straightforward.

The first is the simple fact that the language of necessity and sufficiency evokes absolute standards, for example, the idea that a single disconfirming case refutes a claim of sufficiency or necessity. Social scientists work with imperfect evidence drawn in imperfect ways from a social world that not only is very complex, but that is also subject to a great deal of chance and randomness. The play of the weather alone on what people do and how they feel is enormous and unpredictable. Thus, the very nature of social phenomena makes social scientists shy away from the absolute standards that are an integral part of most discussions of causal necessity and sufficiency. In this chapter, I have shown various ways to relax these standards and bring necessity and sufficiency back into social science discourse (see also Dion 1998).

The second reason has less to do with the nature of social phenomena and more to do with conventional practices in variable-oriented research. Generally, when variable-oriented researchers address causal complexity, they do so by examining the many single causes that they believe may affect some outcome. The implicit model they use is that every single cause is a sufficient cause.[7] In this approach, the goal is to identify all the relevant single causes and then assess their relative importance. Consider the cross-case analysis of strikes one last time. The conventional approach would be to elaborate a list of possible causes (among them, wage levels, unemployment rates, inflation, and union strength), and then assess the relative importance as each cause as an "independent" cause of strikes. The key question would be, Does each cause significantly increase the odds of a strike, holding the others constant, and if so, how much?

Without going into the details of statistical procedures, a researcher using these techniques might find that the introduction of new technology opposed by workers increases the odds of a strike by 5 or 10 per-

7. While not a central concern here, it is worth noting that in the conventional approach causal necessity and sufficiency are approximated, probabilistically speaking, by the model or prediction equation *considered as a whole*, assuming this model has been completely and correctly specified.

cent, the exact figure does not matter, controlling for the impact of other relevant conditions (e.g., wage levels, unemployment rates, union strength). The effects of these other conditions are "held constant" using statistical techniques designed to calculate the "independent" effect of each causal variable.[8] For example, if the introduction of new technology opposed by workers often occurs when there is poor communication between workers and managers, and if poor communication, by itself, also has a positive effect on the odds of a strike, then the estimate of the impact of new technology on the odds of a strike must be downwardly adjusted to take account of its frequent association with poor communication. Conventional statistical procedures are designed to make such adjustments with great precision.

Researchers using conventional techniques almost never examine causal combinations. Their emphasis, instead, is on individual causes and assessing their relative strength. Thus, it is usually difficult to incorporate the consideration of necessity and sufficiency into these analyses because the examination of these aspects of causation usually entails the analysis of combinations of conditions, as sketched in this chapter. In the conventional statistical approach every cause is thought to be capable of acting alone to influence the outcome (i.e., the common bias is toward causal additivity), and every causal variable is usually thought to be causally relevant in the same way for all cases (i.e., the usual bias is also toward causal homogeneity). Thus, researchers using conventional quantitative techniques make very strong assumptions about the nature of the causes they study. They also make very strong assumptions about the homogeneity of the cases that are embraced by their populations.

These practices are antithetical to the study of social diversity, especially as manifested in causal complexity. As I show in this chapter, it is possible to allow causation to be complex, to avoid making assumptions about its specific character, and still study it in a systematic way.

8. In the conventional approach, the assessment of a cause's impact basically involves comparing the odds of the outcome in the first column of table 4.2 (the frequency of cases in cell 1 is divided by the frequency of cases in cell 3) with the odds of the outcome in the second column of table 4.2 (the frequency of cases in cell 2 is divided by the frequency of cases in cell 4) after adjusting for the impact of other causes on the difference in the odds of the outcome. Thus, the conventional statistical assessment of the impact of a cause on an outcome involves all four cells of the simple cross-tabulation presented in table 4.2.

LOOKING BACK, LOOKING AHEAD

Chapters 1–4 establish the basic elements of the diversity-oriented approach, mostly by contrasting this approach with variable-oriented and case-oriented strategies. As I demonstrate in these chapters, the diversity-oriented approach does the following:

1. It treats the constitution of populations as an ongoing process laden with potential for obscuring diversity.
2. It sees cases as complex configurations and recognizes the possibility that a single difference between two cases may signal a difference in kind.
3. It allows for maximum causal complexity in social processes, especially for the possibility that no single causal condition may be either necessary or sufficient for an outcome.

Configurational thinking is at the core of the diversity-oriented approach. Central to configurational thinking, in turn, is the notion of the property space, an analytic device showing all combinations of relevant conditions. In Lazarsfeld's approach to property spaces, the key concern is the elaboration and reduction of types. As I show in these chapters, the property space notion can be extended to the analysis of causal conditions, where it can be used to guide and structure the examination of causal complexity, especially as manifested in causal conditions that are sufficient only in combinations.

The usual lesson that social scientists draw from the observation that no single cause is either necessary and sufficient in the study of social phenomena is that necessity and sufficiency have no place in social research. This chapter demonstrates otherwise, first by showing the importance of studying causation that is both conjunctural and multiple and then by showing how to do so. Chapter 5 provides a detailed example of the study of causal complexity, based on the techniques I present in this chapter. I present two analyses of the same evidence, one using probabilistic criteria, as just discussed, and the other using the "no cell 4 cases" rule (i.e., veristic criteria).

FIVE

THE LOGIC OF DIVERSITY-ORIENTED RESEARCH

INTRODUCTION

This chapter describes a data analytic strategy known as *qualitative comparative analysis* (QCA; see Ragin 1987; Drass and Ragin 1992), originally developed as a formalization and extension of the comparative case-study approach. In a nutshell, QCA provides analytic tools for comparing cases as configurations of set memberships and for elucidating their patterned similarities and differences.[1] With QCA it is possible to view cases as configurations, examine complex patterns of causation, and reconstitute populations based on the patterns that cases exhibit (Ragin 1995). QCA frees social scientists from many of the restrictive, homogenizing assumptions of conventional variable-oriented research without giving up the possibility of formulating statements about broad, cross-case patterns. This alternative approach to cross-case analysis provides the set-theoretic foundation for diversity-oriented research.[2]

Most of this chapter is devoted to describing the application of QCA to *dichotomous* social data reflecting the memberships of cases in conventional, crisp sets. In contrast to statistical methodology, which is based on linear algebra, QCA is based on Boolean algebra, the algebra of sets and logic. QCA treats social categories as sets and views cases in terms of their multiple set memberships. In Boolean algebra a case is either in or out of a set and QCA uses binary-coded data, with 1 indicating membership and 0 indicating nonmembership. Combinations of crisp-set memberships, represented as arrays of binary data, are compared and contrasted to identify decisive cross-case patterns.

Because there is no allowance in QCA for the partial membership of cases in sets, QCA does not directly address one of the main objectives of this book, namely, the introduction of fuzzy sets to the social sciences and the development of fuzzy-set techniques for the analysis of social data. As I detail in part 2 of this work, with fuzzy sets the membership of cases in sets can be partial, with membership scores ranging from 0 (nonmembership) to 1 (full membership). (For example, an Eastern European country might have a membership score of .68 in the set of "rich countries.") This scheme offers a major advance over Boolean algebra and crisp sets.

Boolean algebra, the foundation of QCA, may appear inadequate from the perspective of fuzzy sets—especially given the ubiquity of partial membership in sets. Recall from the introduction, however, that it is impossible for social scientists to take full advantage of the power of fuzzy sets without also adopting techniques of data analysis that focus on set-theoretic relationships. Most quantitative techniques of data analysis eschew the examination of such relationships. However, the analysis of set-theoretic relationships is the core of QCA. Thus, the discussion of QCA offered in this chapter is essentially the presentation of a data analytic strategy that provides a basis for the integration of fuzzy sets into social research. It is important to understand and appreciate the set-theoretic principles that are central to QCA in order to understand how to take full advantage of the power of fuzzy sets. The analytic procedures that I present in part 2 using fuzzy sets parallel those discussed in this chapter.

1. Only some of the procedures described in this chapter are implemented in QCA 3.0 (Drass and Ragin 1992). All are implemented in FS/QCA 1.0 (Drass and Ragin 1999). Procedurally speaking, QCA 3.0 is a subset of FS/QCA 1.0. The new program separates the analysis of necessity and sufficiency, incorporates the use of probabilistic criteria, and fully implements all fuzzy-set procedures. Both QCA 3.0 and FS/QCA 1.0 can be downloaded from the Web site http://www.nwu.edu/sociology.

2. The exact operation of QCA, as described in this chapter, is *not* identical to the description presented in *The Comparative Method* (Ragin 1987). I have altered the procedure slightly so that the parallel to fuzzy-set techniques, presented in part 2 of this work, is enhanced. Specifically, the approach presented in this chapter considers not only all logically possible combinations of conditions—the emphasis of *The Comparative Method*—but also all possible "groupings" of cases. Also, rather than use what is known as a "prime implicant chart" to logically simplify results, in this chapter I show how to use the "containment" rule. Finally, this chapter demonstrates how to incorporate probabilistic criteria into the analysis of set-theoretic relationships, building on the discussion presented in chapter 4.

WORKING WITH CASES AS CONFIGURATIONS

In QCA cases are examined in terms of their multiple memberships in sets, viewed as configurations. This interest in how different aspects or features combine in each case is consistent with an emphasis on understanding aspects of cases in context. For example, having many small- to medium-sized political parties (party "fractionalization") signifies different things about a country's political stability, depending on, among other things, the nature of its electoral system, its degree of sociocultural diversity, and the age of its political institutions. Take another example: having many debts can signal different things about a person's financial situation, depending on his or her other attributes, including age, income, employment status, assets, and career trajectory. By looking at combinations of aspects, it is possible to get a sense of a case as a whole, especially how its different aspects connect. As noted in chapter3, this emphasis on how characteristics connect contrasts sharply with the variable-oriented view of aspects of cases as analytically separable, independent features.

In every social scientific investigation, the selection of cases and attributes to study is dependent on the substantive and theoretical interests of the researcher and his or her intended audiences. Sometimes a research literature is especially well developed, and the selection of cases and attributes is relatively unproblematic. In other situations, however, the researcher can formulate a worthwhile selection of attributes only through an in-depth analysis of cases (Amenta and Poulsen 1994). Sometimes researchers must constitute relevant cases and their key aspects through a systematic dialogue between ideas and evidence, as described in chapter 2. Researchers progressively refine their understanding of relevant cases and their key aspects as they sharpen the concepts appropriate for studying them.

Often the selection of aspects is shaped by the nature of the outcome to be investigated and the researcher's understanding of the causal conditions relevant to this outcome. The selection of causal conditions is usually broad because the concern is to identify not only the factors that seem connected to the outcome as proximate causes, but also the conditions that provide the contexts for the operation of these proximate causes. A fractionalized party system, for example, could be a proximate cause of political breakdown in some situations; in other situations, it might be irrelevant; still in others, it might contribute to long-term political stability. Thus, it is important to consider the contexts and conditions that enable and disable causal connections. This concern for the impact of context on causal connections is a key aspect of configurational thinking.

In QCA, once a set of causally relevant aspects has been identified, the researcher constructs a table listing the different logically possible combinations of these attributes ("configurations") along with the cases that conform to each configuration. As explained in chapter 3, this table can be seen as a "property space" (Lazarsfeld 1937). Each location within a property space, in turn, can be seen, potentially at least, as a different kind or type of case. In QCA, attributes are represented with presence/absence dichotomies, with 0 indicating absence (the case is not in the set in question), and 1 indicating presence (the case is in the set in question). Multichotomies (e.g., race/ethnicity at the individual level) are represented with multiple dichotomies, which can be coded in a variety of ways, depending on the interests of the investigator.

By examining the cases that conform to each configuration, represented as a row of the table, it is possible for the investigator to evaluate whether the best set of attributes has been identified. For each configuration, the researcher asks, Do these cases go together? Are they comparable instances, in the context of the present investigation? Thus, the configurational understanding of cases focuses on the equivalence of cases at the level of the configuration, not simply at the more global, population-wide level.

Consider, for example, table 5.1, which shows different configurations of conditions relevant to ethnic political mobilization among territorially based linguistic minorities in Western Europe. Four attributes define the property space: (1) whether the minority is large or small, (2) whether the minority has a weak or strong linguistic base, (3) whether the minority region is richer or poorer than the core region of the host country, and (4) whether the minority region is growing or declining (see Allardt 1979; Ragin 1987). There are 16 logically possible combinations ("configurations") of these four presence/absence dichotomies, and thus 16 mutually exclusive subsets of cases. For notational convenience in the discussion that follows, the presence of an attribute is denoted by the name of the attribute; the absence of the attribute (negation) is denoted with the "~" symbol preceding the attribute name. Thus, "large" indicates that the linguistic minority is large in size, while "~large" indicates that it is small; "fluent" indicates good linguistic ability; "afluent" indicates poor linguistic ability; "~wealthy" indicates that the minority region is wealthier than the core region; "~wealthy" indicates that it is poorer than the core region; "growing"

Table 5.1
Territorially Based Linguistic Minorities in Western Europe

Combination No.	Size[a]	Linguistic Ability[b]	Relative Wealth[c]	Growth[d]	Instance
1	~large	~fluent	~wealthy	~growing	Lapps, Finland
					Lapps, Sweden
					Lapps, Norway
					Finns, Sweden
					Albanian, Italy
					Greeks, Italy
2	~large	~fluent	~wealthy	growing	North Frisians, Germany
					Danes, Germany
					Basques, Frances
3	~large	~fluent	wealthy	~growing	Ladins, Italy
4	~large	~fluent	wealthy	growing	None
5	~large	fluent	~wealthy	~growing	Magyars, Austria
					Croats, Austria
					Slovenes, Austria
					Greenlanders, Denmark
6	~large	fluent	~wealthy	growing	None
7[e]	~large	fluent	wealthy	~growing	Aalanders, Finland
8[e]	~large	fluent	wealthy	growing	Slovenes, Italy
					Valdotians, Italy
9	large	~fluent	~wealthy	~growing	Sards, Italy
					Galicians, Spain
10[e]	large	~fluent	~wealthy	growing	West Frisians, Neth.
					Catalans, France
					Occitans, France
					Welsh, Great Britain
					Bretons, France
					Corsicans, France
11	large	~fluent	wealthy	~growing	None
12[e]	large	~fluent	wealthy	growing	Friulians, Italy
					Occitans, Italy
					Basques, Spain
					Catalans, Spain
13	large	fluent	~wealthy	~growing	Flemings, France
14[e]	large	fluent	~wealthy	growing	Walloons, Belgium
15[e]	large	fluent	wealthy	~growing	Swedes, Finland
					South Tyroleans, Italy
16[e]	large	fluent	wealthy	growing	Alsatians, France
					Germans, Belgium
					Flemings, Belgium

Note: The (~) sign preceding an attribute name indicates "not" or negation.
[a] Whether the minority is large or small (~large).
[b] Whether the minority has a strong (fluent) or weak (~fluent) linguistic ability.
[c] Whether the minority region is richer (wealthy) or poorer (~wealthy) than the core region of the country.
[d] Whether the minority region is growing or declining (~growing).
[e] Row with strong evidence of ethnic political mobilization.

indicates that the region is growing; "~growing" indicates that it is not growing. Superscripted "e" next to configuration numbers indicates minorities that have mobilized politically (configuration nos. 7, 8, 10, 12, 14, 15, and 16).

Table 5.1 also shows the cases conforming to each of the 16 logically possible combinations of these four dichotomies. By evaluating the comparability of the cases conforming to each configuration, the researcher can make a preliminary assessment of the adequacy of the aspects selected for investigation. For example, the first configuration (~large·~fluent·~wealthy·~growing; note that midlevel dots are used to indicate combinations of characteristics) brings together Lapps in Finland, Lapps in Sweden, Lapps in Norway, Torne Valley Finns in Sweden, Albanians in Italy, and Greeks in Italy. Viewing these six cases together, the researcher asks whether it is reasonable to group these as similar cases in a study of ethnic political mobilization of linguistic minorities in Western Europe. If not, then additional attributes should be added to the list of relevant causal conditions, or perhaps the researcher should substitute different attributes for some of the listed attributes. For example, the investigator may believe that the four minorities in Scandinavia differ in some causally decisive way from the two minorities in Italy. If so, the causal condition that distinguishes these two groups should be added to the table. The cases conforming to each configuration in the property space (i.e., row of the table) should be evaluated in this manner.

When researchers view their evidence in terms of logically possible combinations of conditions along with the cases conforming to each configuration, as in table 5.1, they also evaluate the cases in each row to see if they all display the same outcome or at least roughly comparable outcomes. For example, the researcher would ask: Are the six cases in the first row similar with respect to their ethnic political mobilization? Each row is examined in this manner, so that the researcher can gain some confidence that a viable specification of relevant causal conditions, which comprise the property space for the outcome, has been realized. Obviously, if the cases in a row display widely divergent outcomes or if they are evenly split between contrasting outcomes, the researcher will examine these cases closely and reformulate his or her specification of causal conditions accordingly. This evaluation of cases with respect to outcomes is separate from the first evaluation, just described, where the researcher asks simply whether the cases grouped within each combination of attributes belong together as comparable cases regardless of their outcomes.

When making assessments of outcomes, it is unrealistic to expect that all the cases in each row will be perfectly consistent with respect to the outcome in question. It is very difficult to capture all causally relevant conditions in broad, cross-case analyses. Furthermore, mistakes can be made when assigning cases to sets or when evaluating the evidence with respect to the outcome in question. As an illustration of the general problem, consider the twelfth row of table 5.1. The Friulian and Occitan minorities of Italy occupy this row, along with the Basque and Catalan minorities of Spain. But the two minorities in Spain are more politically active than the two in Italy, and the Basque minority is more mobilized than the Catalan. This information could be used to identify a fifth causal condition, or perhaps even to reformulate the property space altogether, with a completely new set of causal conditions. Alternatively, the researcher might decide simply that there is enough evidence of ethnic political activity across the four cases in this row to justify treating them all as instances of ethnic political mobilization. In other words, the researcher might conclude that the discordance is not so great as to motivate any change in the specification of causal conditions.

The larger point is that the examination of outcomes is a central part of constructing a property space and generating configurations, especially when it comes to the selection of causally relevant aspects of cases. This interplay between theory and data analysis leads to a progressive refinement of the understanding of relevant cases and to a more nuanced elaboration of the ideas guiding the research. Again, perfect consistency in outcomes for the cases with the same combination of causal conditions is relatively rare. As demonstrated in chapter 4, however, it is possible to use probabilistic criteria when evaluating the links between causal conditions and outcomes. This tactic partially ameliorates the problem of contradictory outcomes and thus allows for some discordance in outcomes within configurations.

It is important to understand that in QCA the fundamental analytic unit is the configuration, along with the cases conforming to each configuration. Thus, table 5.1 should *not* be viewed as a presentation of four presence/absence dichotomies but rather as a specification of 16 qualitatively distinct conditions—that is, 16 kinds of cases. The principles of configurational thinking discussed in chapter 3 mandate allowance for the possibility that a single difference between two cases may signal a difference in kind. This thinking provides the conceptual basis for constructing and evaluating evidence in terms of logically possible combinations of causes (i.e., as configurations). Thus, the table should

be viewed as a property space with 16 separate locations. Each of the 16 configurations constitutes, potentially at least, a qualitatively distinct constellation. Note that if five dichotomies had been used, there would have been 32 configurations ($2^5 = 32$); 6 dichotomies would yield 64 configurations, and so on. Using dichotomies, the number of logically possible combinations is equal to 2^k, where k is equal to the number of attributes.

The four dichotomies presented in table 5.1 can be presented not only as 16 configurations (the logically possible combinations of the four presence/absence attributes), but also as 80 logically possible *groupings*. The 16 configurations presented in table 5.1 provide basic groupings, that is, groupings that use all four presence/absence dichotomies. They correspond to the 16 cells of the property space that can be constructed from four dichotomies. Additional groupings can be formed by merging configurations that share one or more attributes. For example, the bottom two rows of table 5.1, linguistic minorities that are large·fluent·wealthy·~growing (the penultimate row) and those that are large·fluent·wealthy·growing (the bottom row), share three attributes and thus can be merged to form a larger grouping, minorities that are large·fluent·wealthy. In set terminology, the larger set is formed from the union of its two component sets. Still larger groupings can be formed from the union of more rows as long as the rows that are grouped contain at least one attribute in common. For example, the first eight rows display ~large. Merging these eight rows yields the set of cases that have ~large in common—that is, all the smaller linguistic minorities.

Just as it is possible to calculate the logically possible number of combinations (2^k), it is also possible to calculate the number of logically possible groupings. The formula is $3^k - 1$, where k again is the number of attributes ($3^4 - 1 = 80$). Table 5.2 shows the 80 logically possible groupings of the four dichotomies presented in table 5.1. Using the formula just described, the 80 possible groupings are formed as follows: 16 involve combinations of four attributes, 32 involve combinations of three attributes, 24 involve combinations of two attributes, and eight involve single attributes.

These 80 groupings are important because they are relevant to any conclusions about cross-case patterns the researcher may wish to construct using the property space presented in table 5.1. For example, the researcher might examine all minorities that are wealthy and growing (wealthy·growing) to see if they all diplay ethnic political mobilization. The 80 groupings listed in table 5.2 provide the basis for formulating

Table 5.2
Groupings Using Four Dichotomies

Initial Configurations (16 Combinations of Four Aspects) (1)	Groupings Involving Combinations of Three Aspects (32) (2)	Groupings Involving Combinations of Two Aspects (24) (3)	Groupings Involving a Single Aspect (8) (4)
~large·~fluent·~wealthy·~growing	~large·~fluent·~wealthy	~large·~fluent	~large
~large·~fluent·~wealthy·growing	~large·~fluent·wealthy	~large·fluent	large
~large·~fluent·wealthy·~growing	~large·fluent·~wealthy	large·~fluent	~fluent
~large·~fluent·wealthy·growing	~large·fluent·wealthy	large·fluent	fluent
~large·fluent·~wealthy·~growing	large·~fluent·~wealthy	~large·~wealthy	~wealthy
~large·fluent·~wealthy·growing	large·~fluent·wealthy	~large·wealthy	wealthy
~large·fluent·wealthy·~growing	large·fluent·~wealthy	large·~wealthy	~growing
~large·fluent·wealthy·growing	large·fluent·wealthy	large·wealthy	growing
large·~fluent·~wealthy·~growing	~large·~fluent·~growing	~large·~growing	
large·~fluent·~wealthy·growing	~large·~fluent·growing	~large·growing	
large·~fluent·wealthy·~growing	~large·fluent·~growing	large·~growing	
large·~fluent·wealthy·growing	~large·fluent·growing	large·growing	
large·fluent·~wealthy·~growing	large·~fluent·~growing	~fluent·~wealthy	
large·fluent·~wealthy·growing	large·~fluent·growing	~fluent·wealthy	
large·fluent·wealthy·~growing	large·fluent·~growing	fluent·~wealthy	
large·fluent·wealthy·growing	large·fluent·growing	fluent·wealthy	
	~large·~wealthy·~growing	~fluent·~growing	
	~large·~wealthy·growing	~fluent·growing	
	~large·wealthy·~growing	fluent·~growing	
	~large·wealthy·growing	fluent·growing	
	large·~wealthy·~growing	~wealthy·~growing	
	large·~wealthy·growing	~wealthy·growing	
	large·wealthy·~growing	wealthy·~growing	
	large·wealthy·growing	wealthy·growing	
	~fluent·~wealthy·~growing		
	~fluent·~wealthy·growing		
	~fluent·wealthy·~growing		
	~fluent·wealthy·growing		
	fluent·~wealthy·~growing		
	fluent·~wealthy·growing		
	fluent·wealthy·~growing		
	fluent·wealthy·growing		

Note: From table 5.1.

any statement that can be made regarding cross-case patterns using this property space. As I show subsequently, the examination of these different groupings is central to the assessment of causal complexity, especially the evaluation of the *sufficiency* of different combinations of causal conditions.

Note also that the 80 groupings in table 5.2 represent all logically possible *selections* on the four causal conditions used in this analysis. Recall from chapter 4 that to assess the sufficiency of a combination of causal conditions, the researcher selects cases with a given combination of conditions and then evaluates whether these cases display the same or roughly the same outcome. If they all (or virtually all) display the outcome in question, then the evidence supports the argument that the combination of causal conditions in question is sufficient for the outcome. The $3^k - 1$ logically possible selections, in this light, can be seen as an attempt to implement an exhaustive examination of causal sufficiency. Each grouping constitutes a different logically possible selection, and each grouping can be evaluated with respect to the outcome. In short, the groupings specify multiple selections on the relevant causal conditions and thus multiple tests of sufficiency.

ANALYZING CAUSAL COMPLEXITY

Usually, social research begins with the goal of explaining an outcome. For example, a researcher might ask why some territorially based linguistic minorities participate in politics on an ethnic basis while others do not. Table 5.1, for example, shows that linguistic minorities in row nos. 7, 8, 10, 12, 14, 15, and 16 (as indicated by the superscripted "e" attached to the row numbers in the table) offer consistent evidence of ethnic political mobilization, while those in the other rows offer weak or no evidence of such mobilization. How should the researcher describe the key differences between these two sets of minorities (mobilized versus not)? In other words, what combinations of causal conditions are linked to ethnic political mobilization?

Recall that in diversity-oriented research, investigators assume maximum causal complexity. This concern for causal complexity is best implemented by allowing for the possibility that no single causal condition may be either necessary or sufficient for the outcome in question. When no single causal condition is either necessary or sufficient, researchers anticipate finding that different *combinations* of causal conditions are sufficient for the outcome. This emphasis on causal complexity, however, does not preclude the possibility of finding necessary

causes. As noted in chapter 4, it is important in any analysis to first test for necessary conditions before examining sufficiency, especially when there is "limited diversity" (i.e., logically possible combinations of causal conditions lacking empirical instances). Necessary conditions may be obscured if researchers focus *exclusively* on sufficiency.

Assessing Necessity

To evaluate necessity, researchers examine instances of the outcome to see if they share any theoretically or substantively relevant causal conditions (see chapter 4). Table 5.3 lists relevant instances of the outcome—the rows from table 5.1 that contain cases with consistent evidence of ethnic political mobilization (seven rows with a total of 19 mobilized minorities). It is clear from simple inspection of the table

Table 5.3
Analysis of Necessary Conditions Using Instances of Ethnic Political Mobilization

Combination No.[a]	Size[b]	Linguistic Ability[c]	Relative Wealth[d]	Growth[e]	Instance
7	~large	fluent	wealthy	~growing	Aalanders, Finland
8	~large	fluent	wealthy	growing	Slovenes, Italy Valdotians, Italy
10	large	~fluent	~wealthy	growing	West Frisians, Neth. Catalans, France Occitans, France Welsh, Great Britain Bretons, France Corsicans, France
12	large	~fluent	wealthy	growing	Friulians, Italy Occitans, Italy Basques, Spain Catalans, Spain
14	large	fluent	~wealthy	growing	Walloons, Belgium
15	large	fluent	wealthy	~growing	Swedes, Finland South Tyroleans, Italy
16	large	fluent	wealthy	growing	Alsatians, France Germans, Belgium Flemings, Belgium

Note: The (~) sign preceding an attribute name indicates "not" or negation.
[a] From table 5.1, show strong evidence of ethnic political mobilization.
[b] Whether the minority is large or small (~large).
[c] Whether the minority has a strong (fluent) or weak (~fluent) linguistic ability.
[d] Whether the minority region is richer (wealthy) or poorer (~wealthy) than the core region of the country.
[e] Whether the minority region is growing or declining (~growing).

that there is no single cause (and thus no combination of causes) that is uniformly present in all instances of the outcome, ethnic political mobilization. The two that come closest are "large size" and "growing," with 16 out of 19 mobilized minorities exhibiting each condition. Thus, based on simple inspection of this table, the researcher would conclude that there are no necessary conditions present among the four conditions that comprise this property space. This evaluation is based on "veristic" criteria. That is, it is based on the simple query: Is it true that the mobilized minorities *uniformly* exhibit one or more causal conditions?

As noted in chapter 4, it is possible to incorporate probabilistic criteria not only into the evaluation of sufficient conditions, as sketched in that chapter, but also into the evaluation of necessary conditions. For example, a researcher might argue that if significantly greater than 65% of the instances of an outcome exhibit the same causal condition, then that condition is "usually necessary" for the outcome. Using this benchmark (.65) and a significance level .05, however, neither "large size" nor "growing" passes the test of "usually necessary." The probability of observing 16 or more successes in 19 trials, with an underlying probability of success equal to .65, is .0591, which exceeds the .05 significance level. To pass .05 significance with nineteen cases, only two cases may deviate (see table 4.9). Thus, even using probabilistic criteria—and thereby permitting disconfirming cases—it is reasonable to conclude that there are no necessary conditions for ethnic political mobilization present in this property space.

Assessing Sufficiency

To assess the sufficiency of a cause or causal combination, the researcher examines the cases conforming to the cause or the combination and evaluates whether they agree in displaying the outcome in question. For example, the evidence presented for row 10 of table 5.1 (cases conforming to the combination ~large·~fluent·wealthy·growing) indicates that this causal combination may be sufficient for ethnic political mobilization because all six cases with this combination display ethnic political mobilization. Of course, researchers must establish standards for evaluating sufficiency and state them clearly. Is six positive cases and no negative cases enough to establish the sufficiency of a causal combination? What about two positive cases and no negative cases, or only one positive case? In each investigation, the investigator must justify the method used to assess sufficiency based on, among other things, the nature of the evidence, previous research, the state of relevant theoretical and substantive knowledge, and the intended audience for the research.

Because the example presented here, ethnic political mobilization among linguistic minorities, involves a moderate number of cases (36), I present both types of sufficiency tests, veristic and probabilistic (see discussion of table 4.10). I focus first on the probabilistic approach, building on the strategies presented in chapter 4. After presenting the probabilistic approach, I present the veristic approach and then contrast the two general ways of assessing sufficiency.

The core of the probabilistic approach to the assessment of sufficiency is to test the significance of the difference between the observed proportion of positive instances and a benchmark proportion specified by the investigator. As explained in chapter 4, the benchmark proportion is linked to linguistic qualifiers, such as "almost always sufficient" (.80) and "sufficient more often than not" (.50). When the number of cases conforming to a causal combination is modest, 30 or fewer, researchers should use an exact probability test; otherwise, the z test for the difference between two proportions will suffice (Hays 1981: 211–14, 647–51). To conduct either the z test or the exact probability test, the researcher must set a benchmark proportion and a significance level before making these assessments. For example, a researcher might argue that if the proportion of cases displaying the outcome in question is significantly greater than .65, using a .05 significance level (one-tailed test), then the causal combination in question is "usually sufficient" for the outcome.

The sufficiency test is applied not only to the original 16 configurations listed in table 5.1 but also to the remaining 64 groupings listed in columns (2)–(4) of table 5.2. In essence, by applying the test to each of the 80 groupings in table 5.2, the researcher examines all logically possible causal expressions that can be constructed from the four presence/absence dichotomies that make up the property space showing causal conditions relevant to ethnic political mobilization. Alternatively, these groupings can be viewed as all possible "selections" on the causal conditions. Table 5.4 summarizes the 80 tests of sufficiency using a benchmark of .65 and a significance level of .05 (one-tailed test). The first column of Table 5.4 lists the 80 groupings. The second column shows the number of cases conforming to each grouping. The third column shows the proportion of cases with strong evidence of ethnic political mobilization (calculated only for groupings with at least two cases). The last column shows the results of the binomial probability tests. Probability levels .05 or lower in the fourth column indicate

Table 5.4
Sufficiency Tests for the 80 Groupings

Grouping (1)	Case (2)	Proportion (3)	Probability (4)
~large·~fluent·~wealthy·~growing	6	.00	
~large·~fluent·~wealthy·growing	3	.00	
~large·~fluent·wealthy·~growing	1		
~large·~fluent·wealthy·growing	0		
~large·fluent·~wealthy·~growing	4	.00	
~large·fluent·~wealthy·growing	0		
~large·fluent·wealthy·~growing	1		
~large·fluent·wealthy·growing	2	1.00	.4225
large·~fluent·~wealthy·~growing	2	.00	
large·~fluent·~wealthy·growing	6	1.00	.0754
large·~fluent·wealthy·~growing	0		
large·~fluent·wealthy·growing	4	1.00	.1785
large·fluent·~wealthy·~growing	1		
large·fluent·~wealthy·growing	1		
large·fluent·wealthy·~growing	2	1.00	.4225
large·fluent·wealthy·growing	3	1.00	.2746
~large·~fluent·~wealthy	9	.00	
~large·~fluent·wealthy	1		
~large·fluent·~wealthy	4	.00	
~large·fluent·wealthy	3	1.00	.2746
large·~fluent·~wealthy	8	.75	.4278
large·~fluent·wealthy	4	1.00	.1785
large·fluent·~wealthy	2	.50	
large·fluent·wealthy	5	1.00	.1160
~large·~fluent·~growing	7	.00	
~large·~fluent·growing	3	.00	
~large·fluent·~growing	5	.20	
~large·fluent·growing	2	1.00	.4225
large·~fluent·~growing	2	.00	
large·~fluent·growing	10	1.00	.0135
large·fluent·~growing	3	.67	.7182
large·fluent·growing	4	1.00	.1785
~large·~wealthy·~growing	10	.00	
~large·~wealthy·growing	3	.00	
~large·wealthy·~growing	2	.50	
~large·wealthy·growing	2	1.00	.4225
large·~wealthy·~growing	3	.00	
large·~wealthy·growing	7	1.00	.0490
large·wealthy·~growing	2	1.00	
large·wealthy·growing	7	1.00	.0490
~fluent·~wealthy·~growing	8	.00	
~fluent·~wealthy·growing	9	.67	.6089
~fluent·wealthy·~growing	1		
~fluent·wealthy·growing	4	1.00	.1785
fluent·~wealthy·~growing	5	.00	
fluent·~wealthy·growing	1		
fluent·wealthy·~growing	3	1.00	.2746
fluent·wealthy·growing	5	1.00	.1160

Table 5.4 *continued*

Grouping (1)	Case (2)	Proportion (3)	Probability (4)
~large·~fluent	10	.00	
~large·fluent	7	.43	
large·~fluent	12	.83	.1513
large·fluent	7	.86	.2338
~large·~wealthy	13	.00	
~large·wealthy	4	.75	.5630
large·~wealthy	10	.70	.5138
large·wealthy	9	1.00	.0207
~large·~growing	12	.08	
~large·growing	5	.40	
large·~growing	5	.40	
large·growing	14	1.00	.0024
~fluent·~wealthy	17	.35	
~fluent·wealthy	5	.80	.4284
fluent·~wealthy	6	.17	
fluent·wealthy	8	1.00	.0319
~fluent·~growing	9	.00	
~fluent·growing	13	.77	.2783
fluent·~growing	8	.38	
fluent·growing	6	1.00	.0754
~wealthy·~growing	13	.00	
~wealthy·growing	10	.70	.5138
wealthy·~ growing	4	.75	.5630
wealthy·growing	9	1.00	.0207
~large	17	.18	
large	19	.84	.0591
~fluent	22	.45	
fluent	14	.64	
~wealthy	23	.30	
wealthy	13	.92	.0296
~growing	17	.18	
growing	19	.84	.0591

Note: From table 5.2.

which causal combinations pass the sufficiency test. As table 5.4 shows, 8 of the 80 groupings pass the sufficiency test.[3]

3. Of course, from a hypothesis-testing viewpoint, many tests, some involving overlapping sets of cases, have been applied to the same "sample" of cases. From a textbook point of view (e.g., Leamer 1978; von Eye 1990), these tests should be adjusted to take this fact into account. I am not opposed to such adjustments. Recall, however, that these techniques are presented primarily as a way to explore data, to enrich the dialogue between theory and evidence. Thus, the ultimate test of the value of this interrogation of the data is its intellectual return: Does it advance understanding of these cases?

Note first that none of the 16 groupings using all four conditions (the configurations from table 5.1) passes the sufficiency test. For a proportion of 1.0 to be significantly greater than .65, with a one-tailed significance level of .05, a grouping needs to have at least 7 cases (see table 4.9). Because none of the 16 configurations has this many cases (6 is the maximum), none passes the sufficiency test. For 6 positive cases and no negative cases to pass a sufficiency test, either the benchmark must be lowered (e.g., to "sufficient more often than not" or .50), or the significance level must be raised (e.g., to .10 significance). Second, observe that the 8 groupings that pass the sufficiency test all have very high proportions: 7 are 1.0; the eighth is .92. Thus, even though the benchmark proportion is relatively modest ("usually sufficient" or .65), only very high proportions with seven or more cases actually pass the test. This result follows from the use of a relatively stringent significance level for evidence of this type.

The eight groupings that pass the sufficiency test are (1) large·~fluent·growing, (2) large·~wealthy·growing, (3) large·wealthy·growing, (4) large·wealthy, (5) large·growing, (6) fluent·wealthy, (7) wealthy·growing, and (8) wealthy. While it is possible to use minimization algorithms to simplify these eight groupings into a logical equation for ethnic mobilization (see Ragin 1987; Drass and Ragin 1992; Drass and Ragin 1999), it is not necessary to do so in this example because the pattern is straightforward. A logically minimal equation can be derived using the *containment rule*. Some groupings are contained within other groupings and thus are logically redundant. For example, linguistic minorities that are large·wealthy·growing (no. 3) are a subset of minorities that are large·wealthy (no. 4), which in turn are a subset of minorities that are wealthy (no. 8). Thus, the third and fourth groupings are *contained within* grouping no. 8 and thus can be eliminated. Altogether, four groupings (nos. 3, 4, 6, and 7) are contained within no. 8, and three are contained within no. 5 (nos. 1, 2, and 3). These logically redundant groupings can be dropped. Eliminating these groupings yields the following simplified statement of the causal conditions sufficient for ethnic political mobilization (as noted previously, in logical statements addition indicates logical *or*):

large·growing + wealthy → ethnic political mobilization.

Using a probabilistic approach to the assessment of causal sufficiency thus produces a relatively parsimonious statement of the conditions for ethnic political mobilization: Linguistic minorities that are wealthy and those that combine large size and growth are the ones that mobilize.

These results confirm that there are no necessary conditions for ethnic political mobilization. However, being wealthy, relative to the core region of the host country, is "usually" sufficient, by itself, for such mobilization. Using this logical statement as a prediction equation yields only one incorrect assignment: Ladins of Italy are a false positive. According to the equation they should offer strong evidence of ethnic political mobilization, but in fact they do not. This deviating case is very complex. Ladins live in a region of Italy that is populated by another territorial minority, South Tyroleans. Unlike Ladins, South Tyroleans are mobilized politically. While every ethnic situation can be considered unique, the situation of Ladins in Italy is clearly more complex than most.

Just as the probabilistic approach to the assessment of sufficiency entails specification of benchmarks and significance levels, the alternative, veristic approach may involve an evaluation of the strength of the evidence using a frequency threshold. In order to enhance the potential for contrast with the probabilistic approach with its implicit frequency threshold of seven positive cases when there are no negative cases, the example of the veristic approach that follows uses a relatively low frequency threshold: If a grouping has no negative instances of the outcome and two or more positive instances of the outcome, it is judged sufficient for ethnic political mobilization. Applying these criteria to the 80 groupings listed in table 5.3 yields the following 23 groupings that pass sufficiency:

1. ~large·fluent·wealthy·growing,
2. large·~fluent·~wealthy·growing,
3. large·~fluent·wealthy·growing,
4. large·fluent·wealthy·~growing,
5. large·fluent·wealthy·growing,
6. ~large·fluent·wealthy,
7. large·~fluent·wealthy,
8. large·fluent·wealthy,
9. ~large·fluent·growing,
10. large·~fluent·growing,
11. large·fluent·growing,
12. ~large·wealthy·growing,
13. large·~wealthy·growing,
14. large·wealthy·~growing,
15. large·wealthy·growing,
16. ~fluent·wealthy·growing,

17. fluent·wealthy·~growing,
18. fluent·wealthy·growing,
19. large·wealthy,
20. large·growing,
21. fluent·wealthy,
22. fluent·growing, and
23. wealthy·growing.

The containment rule described previously can be applied to this list to simplify these 23 causal combinations into a single logical statement. Alternatively, the minimization algorithm I describe in *The Comparative Method* may be used; the results are the same. Applying either technique results in the following logical statement for the causal combinations linked to ethnic political mobilization:

large·growing + fluent·wealthy → ethnic political mobilization.

In short, the results are very similar, though not identical, to those obtained using the probabilistic approach. The equation states that territorially based linguistic minorities that combine either large size and growth or a strong linguistic base and greater relative wealth are the ones that exhibit substantial ethnic political mobilization. In this equation, it is clear that no single condition is either necessary or sufficient because both terms are combinations formed from different causal conditions. These results duplicate those in *The Comparative Method*, where I present a somewhat different, though compatible, approach.

While not as parsimonious as the results using the probabilistic approach, it is easy to see that the two equations differ precisely because the veristic test does not allow false positives. Thus, Ladins of Italy are not covered by the equation that follows from the application of the veristic test of sufficiency. They are excluded because of their weaker linguistic ability, compared with the positive instances of mobilization.

It is not productive at this point to ask, Which solution is correct? because *correctness* is not intrinsic to analytic techniques. Analytic techniques offer social scientists different ways of describing and representing social phenomena (Ragin 1994a). The two equations just presented are alternate presentations of the same evidence on ethnic political mobilization using different techniques for assessing sufficiency. In social research, different techniques almost always result in different portraits of the phenomenon. Which approach is "best" depends on the criterion applied. For example, if the criterion is "no false positives," then the veristic approach may be best. If the criterion is "makes allowance for

imperfect evidence," then the probabilistic approach may be best. As a general rule, when the number of cases is small and researchers can gain in-depth knowledge of cases, the first criterion is more important. When the number is large, the second criterion is more important.

Ultimately, the question of "correctness" can be addressed only through case-level analysis. For example, the investigator might take a close look at the positive instances of ethnic political mobilization where greater relative wealth seems important as a causal factor and examine whether linguistic ability also seems important in these cases. In addition, the researcher could ask whether weaker linguistic ability seems to be the main factor interfering with the development of strong ethnic political mobilization among Ladins in Italy. More generally, as I stress repeatedly in *The Comparative Method* and elsewhere, representations of this type, where large amounts of evidence are reduced to broad patterns summarized in an equation (or using some other shorthand), must be evaluated in terms of their utility for understanding cases. Broad representations are best viewed as maps or guides to help a researcher through difficult terrain (Becker 1986). They cannot show many details, only the most important. As Charles Tilly (1997:54) would argue, representations of this type "discipline our thinking about . . . complex phenomena in preparation for genuine explanatory efforts" at the case level.

CAUSAL COMPLEXITY AND SIMPLIFYING ASSUMPTIONS

While it might seem that producing an equation that provides a summary statement of broad causal patterns concludes the analysis of causal complexity, it does not. No analysis is complete without an examination of the "simplifying assumptions" embedded in a summary statement. As explained in chapters 3 and 4, simplifying assumptions take advantage of limited diversity. Limited diversity exists when one or more of the logically possible combinations of causal conditions specified in an analysis does not exist empirically—a common situation in the study of naturally occurring social phenomena—or when the number of cases attached to a configuration is too small to permit a sufficiency test. In effect, simplifying assumptions assert that cases with these combinations of conditions would pass sufficiency—if, in fact, enough of them existed to conduct the test. In practical terms, investigators make simplifying assumptions whenever one or more of the initial groupings included in an analysis (e.g., the 16 configurations listed in table 5.1) has too few cases to permit a sufficiency test yet is covered by the equation that follows from the analysis of causal complexity. It is not "wrong" to make simplifying assumptions. However, it is hazardous to make them without assessing their plausibility. Thus, it is important for researchers to identify the simplifying assumptions they have incorporated into summary statements and to evaluate them.

For illustration, consider again the results just presented, using the veristic approach to the assessment of sufficiency:

large·growing + fluent·wealthy → ethnic political mobilization.

The first causal combination brings together 4 of the original 16 configurations:

large·growing = large·~fluent·~wealthy·growing +
large·~fluent·wealthy·growing +
large·fluent·~wealthy·growing +
large·fluent·wealthy·growing.

The second causal combination also brings together 4 of the original 16 configurations:

fluent·wealthy = ~large·fluent·wealthy·~growing +
~large·fluent·wealthy·growing +
large·fluent·wealthy·~growing +
large·fluent·wealthy·growing.

At a minimum, the researcher should be able to point to at least one case that displays both the causal conditions and the outcome for each configuration that is covered by an equation. Inspection of table 5.1 shows that this minimum threshold is met by all the configurations covered by the equation for ethnic political mobilization. Thus, the results presented above do not incorporate simplifying assumptions.

Note, however, that if the researcher were to study the combinations of conditions linked to an absence of ethnic political mobilization, which would involve all the rows not covered by the results for ethnic political mobilization, then he or she would find that simplifying assumptions are incorporated into the results of this analysis.[4] Table 5.1 shows clearly that rows 4, 6, and 11 (configurations not covered by the equation for ethnic political mobilization) lack empirical instances.

4. Assumptions about configurations not covered by an equation (e.g., configurations not covered by the equation for ethnic political mobilization) are not as critical or as important as assumptions made about configurations covered by an equation. Each equation constitutes a logically minimal expression of the causal combinations linked to a specific outcome.

Thus, the results of the analysis of the absence of ethnic political mobilization would incoporate the assumption that if found, instances of these absent causal configurations would provide little or no evidence of ethnic political mobilization.

From the perspective of conventional variable-oriented research, the fact that results may be extrapolated to configurations that lack empirical instances is not a serious problem. In this approach, causal conditions are viewed as "independent variables" that have separate effects on the outcome. All that matters is a variable's net effect, controlling for the effects of its competitors. In a configurational approach, however, the basic analytic unit is not the "variable," but the "configuration" and the cases conforming to each configuration. Recall that the key principle undergirding the configurational approach is the idea that a single difference between two cases may signal a difference in kind. When inclusive groupings are formed (e.g., the grouping of minorities that are large·growing), configurations that embrace cases differing on one or more attributes are joined together as equivalent instances. If it can be shown that the component configurations that make up a grouping display the outcome in question, then this merging is justified because the researcher has demonstrated that the configurations are, in fact, equivalent with respect to the outcome. But when larger groupings involve configurations that lack empirical instances altogether, researchers can include these configurations in the summary statement only if they are willing to make assumptions about what these cases' outcomes would be, if in fact such cases existed.

Once an equation showing sufficient combinations of conditions has been derived, therefore, the researcher should examine the simplifying assumptions it incorporates to see if these assumptions are warranted. This evaluation, in turn, should be based on the investigator's substantive and theoretical understanding of the outcome in question. If the simplifying assumptions are not warranted, then the equation should be reformulated excluding the configurations that involve unwarranted assumptions (Ragin 1995).

THE MULTIPLE CONSTITUTION OF POPULATIONS

One hallmark of the diversity-oriented approach is its allowance for both the constituted nature of social scientific populations and the fragility of these constructions. This concern permeates the discussion of QCA presented so far. For example, rather than seeing the 36 territorially based linguistic minorities as a single population, table 5.1 constructs them as 16 different kinds of minorities. Furthermore, in the construction of these 16 different kinds, each configuration is evaluated with respect to the cases that conform to it and also with respect to the outcomes they exhibit. First the researcher asks of each configuration, Do these cases belong together in this study? If they do not, a new scheme is formulated. Second, the researcher asks, Do these cases all have the same outcome or at least roughly comparable outcomes? If there is too much discordance, new causal conditions may be selected, which in turn provide the basis for generating new configurations. When making these evaluations the researcher continually reassesses and reconstructs the property space that defines "kinds" of cases.

This understanding of populations also permeates the evaluation of sufficiency. Tests of sufficiency are conducted for each configuration and for every possible grouping that can be formed from these configurations. In the end, configurations are joined together as similar only if they pass relevant sufficiency tests. Furthermore, once a summary equation is derived, it is evaluated with respect to simplifying assumptions. The researcher must indicate which configurations covered by the equation lack adequate evidence.

In the empirical analysis presented in this chapter, no discussion of the initial selection of the 36 territorially based linguistic minorities was offered. In effect, this set was taken more or less as fixed. However, it is important to emphasize that this larger population is also open to reconstitution in the course of the research. For example, when evaluating the cases that conform to each logically possible combination of conditions, the researcher might decide that some of the smaller minorities are too small and thus do not belong in the same study with the other minorities. Or the investigator might decide that minorities with very weak linguistic ability do not belong. Alternatively, after examining the 36 cases closely the researcher might broaden the analysis to embrace territorially based minorities defined by regional dialects and not just those defined by language differences.

When evaluating the outcomes associated with each configuration, the researcher might also use this as an opportunity to restrict or enlarge the set of cases included in the study. For example, the researcher might find that mobilization is muted or distinct in some other way when the minority speaks a language that is dominant in a neighboring country (Rothschild 1981), for example, speakers of Swedish in Finland. These minorities might be dropped from the investigation. (Alternatively, the researcher might add this contextual factor to the analysis.) More generally, the detailed, case-level evaluations that go into the con-

struction of a property space and the generation of configurations (as in table 5.1) often prompt the investigator to reconsider the set of cases initially chosen for investigation.

Finally, it is also possible to use summary equations, like the equation for ethnic political mobilization, to reconstitute populations. Essentially, a summary equation shows, in a logically shorthand manner, the different combinations of conditions linked to an outcome. These different combinations provide a basis for describing alternate paths to a given outcome, and cases can be classified according to the paths they travel. For illustration, consider again the results of the analysis using the veristic assessment of sufficiency:

$$\text{large}\cdot\text{growing} + \text{fluent}\cdot\text{wealthy} \rightarrow \text{ethnic political mobilization.}$$

In essence, the equation states that there are two sets of conditions linked to ethnic political mobilization: large size combined with growth, and linguistic strength combined with greater relative wealth. Table 5.5 shows the different linguistic minorities conforming to each combination of conditions. (Three minorities, listed in the third column of table 5.5, conform to both causal combinations.)

While far beyond the scope of this chapter, a researcher might find important differences between the nature of the ethnic political mobilization present in these different sets of cases. In fact, an important way to reinforce the results would be to examine the cases to see if differences in the character or course of ethnic mobilization can be traced to differences in relevant causal conditions. In the end, the researcher might be able to differentiate types of ethnic political mobilization and

Table 5.5
Conformity of Cases to Causal Combinations

Minorities That Are large·growing	Minorities That Are fluent·wealthy	Minorities That Conform to Both
W. Frisians, Neth.	Aalanders, Finland	Alsatians, France
Catalans, France	Slovenes, Italy	Germans, Belgium
Occitans, France	Valdotians, Italy	Flemings, Belgium
Welsh, Great Britain	Swedes, Finland	
Bretons, France	South Tyroleans, Italy	
Corsicans, France		
Friulians, Italy		
Occitans, Italy		
Basques, Spain		
Catalans, Spain		
Walloons, Belgium		

assign cases to types (including mixed types) based on these results. Thus, the results provide a basis for reconstituting cases in terms of broad types, based on their causally relevant features. This construction of broad types dovetails with some of the key analytic concerns of diversity-oriented research.

Note that the decision about whether to include or drop a case from an investigation is essentially a dichotomous choice: A case is either "in" or "out" of the study. As I show in part 2 of this work, however, it is possible to apply fuzzy-membership criteria to this assessment so that researchers are not limited to crisp, "in-or-out" decisions. For example, a researcher could gauge the degree to which the evidence justifies including a collectivity in the set of "territorially based linguistic minorities" and operationalize such assessments as a fuzzy set, with membership scores ranging from 0 to 1. This fuzzy set, in turn, could be manipulated as a causally relevant condition, just like any other fuzzy set (see chapter 7). Thus, as I elaborate in part 2, even the degree to which a case warrants inclusion in a study can be evaluated as a fuzzy set and used in the analysis as a contextual factor.

SUMMARY: USING QCA

There are three distinct phases in the application of QCA to cross-case evidence: (1) selecting cases and constructing the property space that defines kinds of cases (initial configurations), (2) testing the necessity and the sufficiency of causal conditions, and (3) evaluating the results, especially with respect to simplifying assumptions. As already noted, the summary equations that result from the application of QCA should be viewed as part of the larger dialogue between ideas and evidence (see also Ragin 1987:164–71). The real test of any representation of evidence is how well it helps the researcher and his or her audiences understand specific cases or sets of cases. Broad representations of cross-case patterns provide maps that guide and facilitate in-depth investigation; they are not substitutes for this type of investigation. Thus, QCA has an implicit fourth phase involving the application of the results to specific cases, but this phase does not belong to QCA proper.

In many respects the first phase is the most difficult (see also Griffin et al. 1991). The dimensions of the property space (i.e., relevant aspects of cases) must be clarified and refined to see if the resulting scheme sorts cases in a way that makes sense. At the same time, the researcher must study the cases initially chosen for the investigation and evaluate whether the set as a whole has integrity. Dropping or adding cases may

help the researcher refine the property space while at the same time increase the comparability of the cases in the study. The researcher also examines cases conforming to each configuration with respect to the outcomes they display. If cases differ too greatly on the outcome, then either the property space must be reformulated or the relevant set of cases must be reconstituted, or both.

Once the researcher successfully stabilizes the relevant cases and the property space that sorts them into types, then the assessment of causal complexity can proceed. In this phase, a key issue is the definition of sufficiency: How should the tests be structured? The answer to this question is shaped in large part by the nature of the evidence and the criteria that are most important to the investigator. Still, in most analyses, it is probably best to work with several definitions of sufficiency and conduct tests favoring competing criteria. Once these tests are complete, QCA can be used to analyze and simplify the patterns.[5]

In the third phase the summary equation itself is examined, especially with respect to simplifying assumptions. The configurational understanding of social phenomena that undergirds diversity-oriented research mandates careful attention to the assumptions that enter into the production of summary statements. Investigators must use their substantive and theoretical knowledge to evaluate these assumptions. Some assumptions may be accepted and others rejected. In the end, the researcher crafts a representation of evidence that reflects both the diversity of cases included in the investigation and the researcher's knowledge.

5. Fuzzy-Set/Qualitative Comparative Analysis or FS/QCA (Drass and Ragin 1999) is a new version of QCA that implements all the features described in this book, including the fuzzy-set methods. It can be downloaded from <http://www.nwu.edu/sociology>.