

Much debate concerning small-N analysis has centered on the question of whether this research tradition has powerful tools for assessing causality. Yet, recent contributions make it clear that scholars are not in consensus with regard to the more basic issue of which procedures and underlying logic are in fact used in small-N causal assessment. Focusing on the field of comparative-historical analysis, this article attempts to clarify these procedures and logic. Methods associated with three major strategies of small-N causal inference are examined: nominal comparison, ordinal comparison, and within-case analysis. The article argues that the use of these three strategies within particular small-N studies has led scholars to reach radically divergent conclusions about the logic of causal analysis in small-N research. One implication of this argument is that methodologists must sort out the interrelationship between strategies of causal inference before arriving at conclusions about the overall strengths and limitations of small-N analysis.

Strategies of Causal Inference in Small-N Analysis

JAMES MAHONEY

Brown University

Studies that qualitatively analyze a small number of cases (i.e., small-N analyses) have generated significant debate in the social sciences. In the past, controversy often focused on whether small-N analysis is a useful approach to scholarly research, especially when compared with statistical analysis (e.g., Lijphart 1971, 1975; Campbell 1975; Smelser 1976; see also Collier 1993). This debate still animates lively methodological commentaries (e.g., Goldthorpe 1997; Ragin 1997; Rueschemeyer and Stephens 1997), but most analysts now agree there is a place for small-N studies in the scholarly cycle of research, even if they disagree about the prominence and importance of this place. Instead of disputes about the overall usefulness of small-N analysis, contemporary methodological debates often focus on the specific procedures of causal analysis used by

AUTHOR'S NOTE: *I would like to thank Jeffery J. Anderson, Andrew Bennett, David Collier, Ruth Berins Collier, Alexander L. George, Jack A. Goldstone, David R. Meyer, and Deborah J. Yashar for helpful comments and suggestions.*

SOCIOLOGICAL METHODS & RESEARCH, Vol. 28 No. 4, May 2000 387-424
© 2000 Sage Publications, Inc.

small-*N* researchers and the underlying logic of inference that characterizes these procedures. On these basic issues, the literature reveals little consensus.

The current disputation has two major axes. The first concerns the actual techniques of causal inference employed in small-*N* research. While in the past much debate focused on whether small-*N* analysts have powerful tools for inferring causality, current disagreements center on which strategies of causal inference are in fact used by small-*N* analysts. Thus, while the long-standing dispute over the usefulness of Mill's methods continues to receive attention,¹ scholars are now in disagreement over the prior issue of whether these methods are even employed by small-*N* analysts. For example, Katznelson (1997:86) argues that the field of comparative-historical analysis shows "a continuing reliance on John Stuart Mill's method of agreement and method of difference," whereas Goldstone (1997:108) maintains that "Mill's methods *cannot* be and *are not* used by comparative case-study works."

The second axis of debate concerns the relationship between causal inference in small-*N* analysis and causal inference in large-*N* statistical research. The issue here is whether small-*N* and large-*N* researchers share a common logic of inference. Ragin and Zaret (1983:733; see also Skocpol 1984; McKeown 1999; Munck 1998) answer in the negative, arguing that the two research traditions "are neither convergent nor congruent in their . . . logic of analysis." By contrast, King, Keohane, and Verba (1994:4) assert that "the differences between the quantitative and qualitative traditions are only stylistic and are methodologically and substantively unimportant. All good research can be understood—and is indeed best understood—to derive from the same underlying logic of inference."

A related controversy is whether small-*N* analysts understand causation in a deterministic manner or in partial and probabilistic terms. Contemporary critics of small-*N* analysis (e.g., Lieberman 1991, 1994, 1998; Goldthorpe 1997) often assert that these analysts use a deterministic method of inference that is inconsistent with the probabilistic model of causation adopted in statistical research. By contrast, some practitioners of small-*N* research (e.g., Collier and Collier 1991:20, 38-39; Goldstone 1997:116-19; Rueschemeyer and Stephens 1997:69) downplay the differences between small-*N* and statistical

research, arguing that small-*N* analysts may use several different techniques of causal analysis, some of which permit the assessment of probabilistic causation.

This article builds on my previous methodological work (Mahoney 1999) by specifying the different strategies of causal inference used in small-*N* analysis and the underlying logic governing these strategies. Although several scholars (e.g., Skocpol and Somers 1980; Skocpol 1984; Nichols 1986; Lieberman 1991) have characterized small-*N* analysis as entailing only one basic approach to causal inference, I argue that analysts select among, and often combine, three different strategies: nominal comparison, ordinal comparison, and within-case analysis. As Table 1 suggests, these three strategies are distinguished along two dimensions: level of measurement and level of aggregation. Both ordinal and nominal strategies involve comparisons at a highly aggregated, cross-case level but are differentiated in terms of their level of measurement. A nominal strategy entails the use of nominal measurement at a high level of aggregation; an ordinal strategy entails the use of ordinal measurement at a high level of aggregation. A within-case strategy is distinct from these two strategies along the dimension of level of aggregation. Whereas nominal and ordinal strategies involve highly aggregated comparisons across cases, within-case analysis entails a shift toward disaggregation and a focus on comparisons within particular cases. Within-case analysis does not contrast with ordinal and nominal strategies in terms of level of measurement. In fact, as we shall see, within-case analysis can be used in conjunction with multiple levels of measurement, including nominal, ordinal, and interval measurement.

Important contrasts in the logic of causal inference underlie nominal and ordinal strategies. A nominal strategy implicitly or explicitly assumes a deterministic understanding of causation built around the ideas of necessary and sufficient conditions. This understanding of causation is quite different from that employed by most large-*N* researchers. By contrast, an ordinal strategy is more compatible with a probabilistic understanding of causation and, thus, more consistent with the assumptions that guide large-*N* research. The strategy of within-case analysis can be used in conjunction with either a deterministic or a probabilistic understanding of causation. In fact, within-case analysis can involve at least three different procedures: pattern

TABLE 1: Strategies of Causal Inference in Small-*N* Analysis

		<i>Level of Measurement</i>		
		<i>Nominal</i>	<i>Ordinal</i>	<i>Interval</i>
<i>Level of Aggregation</i>	<i>Aggregated</i>	Nominal Strategy	Ordinal Strategy	Not Typically Used
	<i>Disaggregated</i>	Within-Case Strategy		

matching, process tracing, and causal narrative. Each raises interesting questions with regard to the relationship between small-*N* and large-*N* research.

In analyzing these different strategies, I seek to avoid presenting an overly stylized portrait of small-*N* analysis and instead focus on the concrete procedures that analysts actually employ when doing research. To this end, I draw on examples from several published works of comparative-historical analysis. My goal is not to fully specify all of the procedures through which these comparative-historical works make causal inferences but, rather, to use them for illustrative purposes in highlighting the different logics and procedures that characterize small-*N* research.

NOMINAL COMPARISON IN CROSS-CASE ANALYSIS

Small-*N* researchers often pursue causal analysis through nominal comparison across cases. Nominal (or categorical) comparison entails the use of categories that are mutually exclusive (cases cannot be classified in terms of more than one category) and collectively exhaustive (one of the categories applies to each case). With respect to different levels of measurement, nominal categorization is sometimes considered unsophisticated because it does not involve the rank ordering of cases, much less quantifying the degree to which particular cases differ from one another. Yet, for conceptualizing certain kinds of phenomena, nominal categories are highly appropriate (Sartori 1987:182-85; see also Collier and Adcock 1999).

Nominal comparison can be used by small-*N* researchers in several different ways. For example, it can be used for the descriptive purpose of clearly stating or summarizing an overall causal argument. Vivid labels can be used in conjunction with nominal categories, and thus scholars may be drawn to these categories when summarizing the different values of explanatory outcome variables at the beginning and end of their studies. Indeed, highly nuanced, complex small-*N* arguments can often be stated in only a few paragraphs or pages when nominal categories are employed. Likewise, nominal comparison can be used as a preliminary step in causal investigation. In this case, the analyst conceptualizes variables as nominal categories simply to develop an initial sense of whether potential explanations and outcomes are roughly matched. The central analysis pursued by the investigator, however, may rely on alternative methods of causal assessment.

Nominal comparison across cases can also be used as a central strategy of hypothesis assessment, which is the focus here. In this section, I argue that small-*N* analysts often employ nominal comparison to eliminate rival explanations based on the logic of necessary and sufficient causation and have great difficulty using nominal comparison as a strategy for evaluating partial and probabilistic causation. As we shall see below, however, deterministic nominal methods are not the only or necessarily the most important tool of causal inference in small-*N* analysis.

DETERMINISTIC METHODS: ELIMINATING POTENTIAL NECESSARY AND SUFFICIENT CAUSES

Scholars frequently describe small-*N* explanations as deterministic. The concept of determinism, however, is often not clearly defined. For some scholars, determinism refers to a quality of the world that must be explicitly theorized. In this formulation, a deterministic model assumes that causal patterns in a given domain are in principle fully predictable, even if analysts can summarize those patterns only in probabilistic terms. Similarly, a nondeterministic model refers to a domain of cases in which inherently stochastic properties make full prediction impossible even in principle. Unfortunately, because social scientists have not devised tools for theorizing which parts of the

world are inherently deterministic and inherently stochastic, this definition has little bearing on contemporary research practices.

When small-*N* analysis is described as deterministic, a more specific definition of determinism is used. In this alternative conception, which I employ here, determinism refers to the way in which explanatory variables are assumed to affect outcome variables (see Ragin and Zaret 1983:744; Lieberson 1994:1228; Skocpol 1984:378). In particular, a deterministic explanation is one in which explanatory variables (or combinations of explanatory variables) are treated as potential necessary and/or sufficient causes of an outcome (Sobel 1995:5).

A deterministic approach to explanation sometimes involves the effort to identify invariant patterns of association. Thus, if a given factor or combination of factors is understood to be the necessary *and* sufficient cause of an outcome, the outcome will always be present when the cause is present. Similarly, when the necessary and sufficient cause is not present, the outcome will always be absent. Yet, a causal factor or combination of factors that is *either* necessary *or* sufficient, but not both, will not follow a pattern of invariant association. When a sufficient cause is present, the outcome will always also be present. However, if a sufficient cause is absent, the outcome could be either present or absent. If a necessary cause is absent, the outcome will always be absent. However, if a necessary cause is present, the outcome could be either present or absent. Hence, a deterministic explanation need not reveal an invariant association between cause and outcome across all possible scores on the explanatory variable.

Several of the major nominal techniques used by small-*N* comparativists are deterministic in this sense of necessary and sufficient conditions, including, most notably, J. S. Mill's method of agreement² and method of difference.³ Although these Millian methods do not permit the analysis of multiple explanatory factors or interaction effects, they provide a sound logical basis for *eliminating* potential necessary and sufficient causes. Specifically, the method of agreement can be used to eliminate potential *necessary* causes, whereas the method of difference can be used to eliminate potential *sufficient* causes. Thus, with the method of agreement, the outcome of interest is present in all cases. Consequently, it is logically impossible for any hypothesized cause that is not shared by the cases to be necessary for the outcome's occurrence, since some cases possess the outcome but not the cause.

By contrast, with the method of difference, the outcome is present in some cases and not present in others. Hence, any hypothesized cause that is shared by all the cases cannot be a sufficient cause of the outcome, since not all cases with the hypothesized cause experience the outcome of interest.

Comparative-historical researchers frequently use these methods to eliminate rival hypotheses. Good examples are found in the sociological literature on the causes of revolutions. For instance, Skocpol (1979:34, 113) uses Mill's methods of agreement and difference to eliminate relative deprivation, system disequilibria, multiple sovereignty, transformative ideologies, and urban worker revolts as causes of social revolutions in France, Russia, and China. Goodwin (2000; see also Goodwin and Skocpol 1989) employs the method of difference to eliminate poverty, social misery, and professional revolutionary organizations as causes of Third World revolutions. And even though Goldstone (1997:108) asserts that small-*N* analysts do not use Millian methods, he implicitly uses these methods himself to reject demographic explanations of the Meiji Restoration, sociological explanations of the French Revolution, and Whig and revisionist explanations of the English Revolution (see Goldstone 1991:405-6, 67, 146).

Scholars also use nominal comparison to demonstrate that their own favored explanations *cannot* be eliminated under the assumptions of deterministic causation. Small-*N* researchers often frame their central causal analyses around a discussion that shows how their explanation is the only one among several considered that cannot be eliminated using nominal-comparative techniques (e.g., Downing 1992; Ertman 1997; Esping-Anderson 1990; Fulbrook 1983; Gould 1999; Orloff 1993; Skocpol 1992; Yashar 1997). This practice gives many small-*N* arguments a compelling underlying logic, especially if the favored explanation is situated within an interesting, broader theoretical tradition.

Mill's methods of agreement and difference are not the only methods that assume a deterministic understanding of causation. For example, the "most similar systems design" and the "most different systems design" have the same logical structure as the method of difference and the method of agreement, respectively, except that Przeworski and Teune's (1970) original formulation of these designs is based on the

premise that the scholar is combining two levels of analysis. Boolean algebra is distinct from these other methods because it allows the analyst to treat several different *combinations* of variables (i.e., causal expressions) as the causes of an outcome. However, as Ragin (1987:100; see also Griffin and Ragin 1994:9; Amenta and Poulsen 1994:23-24) has suggested, Boolean algebra is "highly compatible" with the idea of necessary and sufficient conditions, and "the results of Boolean analysis are easy to interpret in terms of necessity and sufficiency."⁴

A nice illustration of Boolean analysis is Wickham-Crowley's (1992) work on the origins of peasant support for guerrilla movements in Latin America. Wickham-Crowley focuses on 20 cases and looks at four different explanatory variables: agrarian structure (A), agrarian disruption (B), rebellious cultures (C), and peasant linkages (D). He first pools cases in which the outcome of peasant support for guerrillas is present and identifies all combinations of scores on explanatory variables that are associated with this outcome. Next, he reduces the number of these combinations by assuming that if two combinations of explanatory variables differ in their scores for a single variable, then that variable can be eliminated from the combination. The implicit rationale behind this Boolean reduction procedure is that the variable is not *necessary* for the combination to have a causal effect, since it is both present and absent in combinations associated with the outcome. For example, all four of Wickham-Crowley's explanatory variables are present in one combination with peasant support for guerrillas (expressed as ABCD), whereas the explanatory variables A, B, and C, but not variable D, are present in another combination with peasant support for guerrillas (this is expressed as ABCd). Hence, Wickham-Crowley assumes variable D is irrelevant to the combination and eliminates it, reducing the causal expression to ABC.

Through this Boolean reduction procedure, Wickham-Crowley (1992) narrows the range of explanations down to four possible combinations of variables under which peasants have supported guerrillas: ABD, AC, CD, and abD. Although Wickham-Crowley does not identify any *individual* causal factor that is necessary or sufficient to produce strong support for guerrillas, the final four expressions are each understood to represent a *combination of factors* that are *sufficient* for peasant support. For example, the causal combination of ABD is

understood to be sufficient for strong support for guerrillas in modern Latin America (i.e., when the combination is present, the outcome is always also present; when it is absent, the outcome may be either present or absent). In this sense, a deterministic understanding of causation characterizes small-*N* analyses that employ Boolean algebra.

LIMITATIONS OF NOMINAL METHODS

Nominal methods are not flawless tools for identifying necessary and sufficient conditions. For one thing, it is always possible that explanatory variables not considered in the analysis might avoid elimination if they were included. In addition, variables eliminated in the Boolean reduction procedure might *not* have been eliminated if other variables had been introduced in the analysis. This is a standard problem of specifying an explanatory model that arises in many forms of causal assessment, and its consequence for nominal methods is that investigators can never know for certain whether they have correctly identified all necessary and/or sufficient variables or combinations of variables. Furthermore, nominal methods cannot be used to weigh the relative importance of explanatory variables or combinations of variables that are not eliminated. The analyst must rely on other criteria to determine which necessary and/or sufficient causes are especially deserving of analytic attention.

Finally, if additional cases were considered, it is possible that presumed necessary and/or sufficient causes would be eliminated. This suggests that the particular selection of cases may strongly influence findings, such that the small-*N* researcher's results may not be consistent with the finding that would emerge if a more representative sample had been considered (Geddes 1990). However, in many domains of small-*N* research, the definition of the population is not clear. Hence, the idea of a representative sample vis-à-vis an overall population may not always be an appropriate standard for evaluating this type of research (Collier and Mahoney 1996).

Moreover, if a clear definition of the population can be established, only a small number of cases are often needed to meet a high level of statistical confidence when employing a deterministic understanding of causation. Indeed, as Dion (1998) has shown, only five cases will usually be enough to yield 95 percent confidence about necessary

causes. Hence, scholars might rethink how appropriate the idea of a “degrees of freedom problem” really is for small-*N* analysts who adopt an understanding of causality built around the ideas of necessary and sufficient causes.⁵ When this understanding of causality is adopted, a small number of cases will often meet the confidence demands of standard statistical analyses.

The usefulness of nominal methods depends in part on whether one believes necessary and sufficient causation is a helpful way to think about causality in small-*N* research. From one perspective, this approach might be seen as unhelpful. Many causal factors are neither necessary nor sufficient, but rather probabilistic, and these methods offer little leverage in evaluating such probabilistic causes. For example, Lieberman’s (1991) well-known account of drunk driving and automobile accidents suggests how the methods of agreement and difference might mistakenly eliminate important causes. Thus, with the method of agreement, the small-*N* analyst who examines three cases of automobile accidents will eliminate drunk driving as a cause if it is present in only two of the three cases. Similarly, with the method of difference, the small-*N* analyst will eliminate drunk driving as a cause of automobile accidents if it is present both in cases of accidents and in cases of nonaccidents. It is essential to recognize that this example does not call into question the ability of nominal methods to evaluate necessary and sufficient causation. These methods correctly show that drunk driving is neither a necessary nor a sufficient condition for an automobile accident (i.e., some automobile accidents occur in the absence of drunk driving, and not all instances of drunk driving produce automobile accidents). Rather, the example suggests the limitations of thinking about causation in terms of necessary and sufficient conditions.

Additional reasons exist for viewing the identification of necessary and sufficient conditions as an unproductive approach to causal inference. Regarding specifically the *necessary* causes of an outcome, there are potentially an infinite number of such causes for any outcome, most of which are unimportant or trivial (e.g., the act of driving a vehicle is a necessary cause of an automobile accident). With regard to the *sufficient* causes of an outcome, many of these causes are obvious or tautological (e.g., intentionally smashing one’s car into another car is a sufficient cause of an automobile accident).

From a different perspective, however, several arguments can be made in defense of the practice of identifying necessary and sufficient causes. First, small-*N* analysts do not randomly and unreflectively select explanatory variables for consideration but, rather, choose them because they are important or interesting in light of the theoretical literature relevant to their research question. Thus, trivial potential necessary causes or tautological potential sufficient causes are not usually considered by small-*N* analysts. Second, many small-*N* analysts focus on *combinations* of variables when using nominal methods. While a particular variable may itself be neither a necessary nor a sufficient cause of an outcome, the variable may be *part of a combination* that is necessary or sufficient for an outcome. Hence, the fact that no single variable is necessary or sufficient for an outcome does not mean that a nominal strategy is of no value. Third, the analysis of necessary and sufficient conditions might be seen as important when evaluating certain outcomes of exceptional interest. For example, many small-*N* researchers believe that information about the necessary or sufficient causes of social revolutions, industrialism, fascism, social welfare, and human genocide can be quite valuable.

Finally, although nominal methods do not permit the assessment of probabilistic causation, small-*N* researchers have other methods available to them that can be used to evaluate probabilistic causal effects. For example, to determine whether drunk driving—a probabilistic cause—plays a role in some automobile accidents, the researcher might employ ordinal analysis or one or more forms of within-case analysis (discussed below). Alternatively, the researcher could explore whether drunk driving *in combination with* other variables is a sufficient cause of automobile accidents in a specific population of cases.

One final concern about the use of nominal methods to assess necessary and sufficient causation should be addressed. Some commentators (see especially Lieberman 1991) have pointed out that these methods assume the analyst is able to measure all variables correctly, since a change in the scoring of a variable for a single case could lead to opposite conclusions about that variable's causal relevance. In this sense, *nominal methods do not permit any notion of measurement error*. However, it is not clear how important a criticism is raised by this observation. Small-*N* analysts are typically experts on each of

their cases, and given that with this method they conceptualize variables as nominal—not continuous—categories, they may indeed be able to avoid measurement error for all of their variables. Moreover, if they do incorrectly score a particular variable for one or more cases, it is likely that other case experts or small-*N* analysts will identify this error, since much debate in this kind of research entails arguments about the scoring of particular variables for specific cases.

NOMINAL METHODS AND PROBABILISTIC EXPLANATION

Some scholars have asserted that nominal methods are compatible with probabilistic explanation in small-*N* research (e.g., Savolainen 1994). This argument may make sense for analysts who use nominal comparison simply to develop an initial sense of which explanatory variables are roughly matched with outcomes before moving on to alternative methods of hypothesis evaluation. However, the real strength of nominal methods rests with their ability to eliminate potential causal factors when only a small number of cases are selected. Small-*N* researchers typically must adjudicate between a large number of rival hypotheses, and causal inference is fundamentally a process of narrowing down the number of potential explanations. When small-*N* analysts do not use nominal methods for the purpose of eliminating rival explanations, they are forfeiting the principal leverage offered by these methods.

Analysts who are convinced that nominal methods can be used to eliminate hypotheses without assuming a deterministic understanding of causation need to specify the logical basis on which such elimination can take place. The problem is that, with only a small number of cases, it is difficult to specify such a basis. Certainly, when assessing a potential explanatory variable under the assumptions of partial and probabilistic causation, the presence of one or two cases that deviate from an overall or expected pattern of matching is not enough to eliminate the variable. For example, if a potential causal variable is present in only three of four cases that share the same outcome, the analyst cannot easily justify eliminating this factor when causal patterns are assumed to be partial and probabilistic: There is no statistical basis for rejecting the factor even if an alternative factor has been identified that is present in all four of the cases. In this sense, scholars who use

nominal methods to eliminate causal factors must at least implicitly assume an understanding of causation as entailing necessary and/or sufficient causes. Otherwise, they must be using one or more alternative strategies of causal inference, not exclusively a nominal strategy.

There is indeed a potential risk in mistakenly believing that nominal methods are compatible with partial causation: Scholars may inadvertently apply their deterministic eliminative logic to explanations they seek to reject but appeal to partial and probabilistic causation when assessing their own explanation. In effect, they may hold alternative explanations to a much higher standard of falsification than their own explanation. Obviously, if researchers are willing to reject other scholars' arguments on the grounds that they cannot meet the standard of deterministic causation, these researchers must also hold their own explanations to this high standard.

ORDINAL COMPARISON IN CROSS-CASE ANALYSIS

In addition to sometimes employing a nominal strategy, small-*N* researchers often conceptualize variables as present to differing degrees across cases, even if they do not usually specify those degrees with numerical coefficients. Thus, ordinal, cross-case comparison is employed by small-*N* analysts. Unlike nominal categorization, ordinal categorization entails rank ordering cases into three or more categories based on the degree to which a given phenomenon is present. This type of comparison is the basis for a second fundamental strategy of causal inference employed in small-*N* research: Mill's method of concomitant variation (Mill [1843] 1974; DeFelice 1986; Mahoney 1999).

METHOD OF CONCOMITANT VARIATION

The method of concomitant variation seeks to establish causation by looking at the relationship between values on an ordinally measured explanatory variable and values on an ordinally measured outcome variable. For example, if values on an explanatory variable and an outcome variable are measured as high, medium, and low, then cases are compared to see if there is an association (possibly an inverse

association) between the two variables. If there appears to be a "strong" association, the analyst may infer the relationship is causal.

This observation raises the issue of what constitutes a strong association for the small-*N* researcher. Mill ([1843] 1974:402-6) argues that, when using the method of concomitant variation, a perfect matching between cause and outcome is not required to infer causality. Hence, whereas one "exception" is enough to eliminate a given causal factor with the methods of agreement and difference, this standard need not apply to the method of concomitant variation.⁶ Furthermore, the method of concomitant variation increases the probability that a less than perfect matching between scores on an explanatory variable and scores on an outcome variable for a small number of cases might reflect causation. A simple example illustrates this point.

Imagine two small-*N* researchers each working with six cases. The first researcher employs the methods of agreement and difference and generates the results listed in Table 2. The second researcher uses the method of concomitant variation and generates the findings presented in Table 3. In both instances, the researcher identifies one case (i.e., case 4) in which scores on the explanatory variable and outcome variable depart from the general pattern of matching found across the full set of cases. For the first researcher (Table 2), who uses the methods of agreement and difference, we may ask: What is the likelihood of matching five out of six cases simply by chance? The answer depends in part on how often the cause is present in the full spectrum of cases that are relevant to the research question,⁷ information that may not be available to the small-*N* researcher. If for the sake of illustration we assume that the distribution of scores on the causal variable within the sample is consistent with the population as a whole, the probability of randomly generating a sample in which at least five out of six cases have scores on the cause and outcome that are matched is nearly 11 percent (i.e., .11).⁸ Given that small-*N* analysts typically consider many potential explanatory variables (often more than 10), finding the pattern in Table 2 for at least one explanatory variable is not unusual and, by itself, provides a weak basis for inferring causality. Because the methods of agreement and difference are specifically designed to *eliminate* causes, it is not even clear how the small-*N* researcher would "explain" the one deviant case.

TABLE 2: Example of Less Than Perfect Matching for the Methods of Agreement and Difference

<i>Case</i>	<i>Cause</i>	<i>Outcome</i>
1	X	Y
2	X	Y
3	X	Y
4 ^a	Not X	Y
5	Not X	Not Y
6	Not X	Not Y

a. Case in which scores on the cause and outcome depart from the general or expected pattern of matching found across the full set of cases.

TABLE 3: Example of Less Than Perfect Matching for the Method of Concomitant Variation: A Case of One "Error"

<i>Case</i>	<i>Cause</i>	<i>Outcome</i>
	(1 = lowest, 4 = highest)	(1 = lowest, 4 = highest)
1	1	1
2	1	1
3	2	2
4 ^a	3	2
5	3	3
6	4	4

a. Case in which scores on the cause and outcome depart from the general or expected pattern of matching found across the full set of cases.

TABLE 4: Example of Less Than Perfect Matching for the Method of Concomitant Variation: A Case of Four "Errors"

<i>Case</i>	<i>Cause</i>	<i>Outcome</i>
	(1 = lowest, 4 = highest)	(1 = lowest, 4 = highest)
1	1	1
2 ^a	1	2
3 ^a	2	1
4 ^a	2	3
5 ^a	3	2
6	4	4

a. Case in which scores on the cause and outcome depart from the general or expected pattern of matching found across the full set of cases.

By contrast, the method of concomitant variation is not a deterministic method, and less than perfect association can be attributed to stochastic processes, measurement error, or additional causal factors. In fact, the association between cause and outcome in Table 3 can be interpreted as quite strong. If we again assume that cases are selected on the dependent variable and the distribution of values on the causal variable in the sample is consistent with the relevant universe of cases, the probability of randomly generating results that covary to the degree in Table 3 is around 1 out of 350 (i.e., .0029).⁹ Even with the small sample of six cases, the likelihood of these results occurring simply because of chance is small enough to believe the relationship may be causal. Thus, the method of concomitant variation not only avoids the deterministic assumptions of other Millian methods but provides some basis for inferring causality when cause and outcome do not reflect a pattern of perfect matching.

However, the primary basis for employing probabilistic criteria is derived from the number of cases selected, and the statistical gain of switching from nominal to ordinal comparison in small-*N* analysis is modest. For instance, Table 4 presents an ambiguous example in which four cases have scores on the causal and outcome variables that depart from an expected pattern of matching. The data do not provide a statistically sound basis for believing the pattern of matching reflects causation (the correlation is not significant at a .05 level). Yet, at the same time, the data do not provide strong grounds for rejecting the causal factor outright (the Spearman rank-order coefficient is .621). This example points to an important limitation of ordinal comparison in small-*N* research: The strategy often does not provide a clear basis for eliminating potential explanatory variables when causal patterns are assumed to be partial and probabilistic. Indeed, to have any confidence that a potential causal variable can be eliminated, the analyst must demonstrate that there is no clear association between values on the causal variable and values on the outcome variable.

THE METHOD OF CONCOMITANT VARIATION IN PRACTICE

The method of concomitant variation has been used by small-*N* researchers both as a supplementary method to support an argument

that is principally developed through nominal comparison and as a central method in its own right for assessing alternative hypotheses. It is instructive to consider examples of both uses here.

Luebbert's (1987, 1991) work on interwar regimes in Europe employs the method of concomitant variation to strengthen a primarily nominal argument. In his overarching nominal assessment, Luebbert argues that the presence or absence of "lib-labism" (i.e., a liberal party-labor alliance) before World War I explains liberal versus nonliberal regime outcomes during the interwar period. Thus, when lib-labism was present before World War I, as in England, Switzerland, and France, a liberal regime developed in the interwar period. By contrast, when lib-labism was absent, as in Norway, Sweden, Denmark, Germany, Italy, and Spain, a nonliberal regime developed. Hence, the dichotomous explanatory variable of lib-labism is perfectly correlated with interwar liberalism.

Although his main explanation relies on nominal comparison, Luebbert reconceptualizes lib-labism and interwar liberalism as variables that are ranked across cases. Table 5 presents this ranking for the main countries considered by Luebbert. If the two ordinal variables are compared across cases, it is apparent that there is substantial, but not perfect, matching. Of the 11 cases, 6 (France, Belgium, Netherlands, Denmark, Spain, and Germany) maintain their rank order, 3 (Switzerland, England, and Norway) move only one rank order, and 2 (Italy and Sweden) move three rank orders. Lib-labism thus emerges as strongly, but not perfectly, associated with liberalism.

The use of ordinal comparison allows Luebbert to state certain findings in a far more nuanced way than is possible when nominal categories are strictly employed. For example, Luebbert points out that among the countries in which lib-labism failed before World War I, Belgium and the Netherlands "most closely approximated the British-French-Swiss pattern of liberal hegemony" (Luebbert 1991:56). That is, these two countries were borderline success cases vis-à-vis the dichotomous explanatory variable of lib-labism. In turn, Belgium and the Netherlands established interwar regimes in which significant liberal elements were present, making them a subtype of liberal regimes (Luebbert 1991:248, 250).

What about cases in which scores on lib-labism and liberalism do not match? In countries that differed only marginally from a perfect

TABLE 5: Method of Concomitant Variation in Luebbert's Analysis of Interwar Regimes

	<i>Prewar Lib-Labism^a</i>	<i>Interwar Liberalism</i>
	(1 = least, 10 = most)	(1 = least, 10 = most)
Switzerland	10	9
England	9	10
France	8	8
Belgium	7	7
Netherlands	7	7
Denmark	6	6
Italy	5	2
Norway	4	5
Spain	3	3
Sweden	2	5
Germany	1	1

SOURCE: Variables are ranked across cases based primarily on the following passages of Luebbert's 1987 ("Social Origins") and 1991 (*Liberalism, Fascism, or Social Democracy*) works: Switzerland (1991:49, 166), England (1987:452; 1991:37-48), France (1987:455-56; 1991:37-48), Belgium and the Netherlands (1987:451; 1991:56, 57, 248, 250), Denmark (1987:451; 1991:57-58, 236-37, 270-71), Sweden and Norway (1991:57-58, 239-42, 270), Italy (1991:57-58, 272-77), Spain (1991:151-53, 272-77), and Germany (1991:272-77).

NOTE: Lib-labism = a liberal party-labor alliance.

a. Luebbert measures lib-labism as "degree of liberal hegemony" present before World War I.

match, such as England and Switzerland, measurement error or an inherently stochastic process may have been at work. Luebbert does not attempt to explain these slight imperfections: His causal analysis is probabilistic in precisely this sense. For the two cases that are more significantly out of line (Sweden and Italy), Luebbert takes steps to identify "missing variables" that might account for the more serious misalignment between cause and outcome. With Italy, for example, Luebbert argues that liberals were unable to consolidate a more significantly liberal regime between the wars—a surprising outcome given their prewar level of lib-labism—because of the extremely fragmented nature of the secular middle-class community, which led to an extremely weak Liberal Party (Luebbert 1991:242-43). Luebbert thus explains the Italian "outlier" by drawing on an additional explanatory factor (i.e., middle-class fragmentation) not part of his original hypothesis.

In her work on social revolutions, Skocpol (1979) also uses ordinal comparison to supplement what is primarily a nominal argument. She

does so by disaggregating nominal variables into constituent subvariables that are evaluated through ordinal comparison (see Mahoney 1999). Hence, while Skocpol's explanatory variable of "conditions for state breakdown" is treated as a dichotomous variable for the purpose of using the methods of agreement and difference, this variable is disaggregated into three constituent subvariables (international pressure, state autonomy, and agrarian backwardness) when evaluated through ordinal comparison (pp. 154-57). These constituent variables are ranked across all cases of revolution and nonrevolution. Likewise, Skocpol's dichotomous outcome variable of social revolution is assessed in terms of ordinal constituent processes. Social revolution is defined in part as "rapid, basic transformations of a society's state and class structures" (p. 4), and Skocpol notes differences between cases along these dimensions. For example, social revolution unfolded most rapidly in Russia, least rapidly in China, and at an intermediate pace in France. These differences are explained in part by ordinal contrasts on key explanatory variables, including the extent of international pressure that marked the revolutionary process (p. 172). In sum, even though Skocpol's book is famous for its use of nominal comparison through the methods of agreement and difference, ordinal analysis plays a major role in underpinning the nominal argument.

While Luebbert and Skocpol use ordinal analysis to strengthen and support their overall nominal arguments, other scholars use ordinal comparison as the principal cross-case method of investigation. One example is Collier and Collier's (1991) work on labor incorporation in eight Latin American countries. Collier and Collier identify four types of labor incorporation periods and seek to explain ordinal differences in the "scope of mobilization" that characterized these periods. Collier and Collier first eliminate certain explanatory factors that lack any consistent relation to the scope of mobilization; for example, they reject explanations centered on the strength of the labor movement because there is "no systematic relationship between labor movement strength and type of incorporation period" (p. 750). Collier and Collier then show how their main explanatory variable—the political strength of the oligarchy—does reveal a clear pattern with mobilization during the incorporation period. In particular, Collier and Collier show there is "an inverse relation between the political strength of the oligarchy . . . and the degree to which . . . mobilization was pursued in

the incorporation period" (p. 748). Although there is a clear inverse relationship for six of the eight cases, two cases deviate from this pattern. In Peru and Argentina, the oligarchy was in many spheres quite powerful, yet these cases exhibited relatively high levels of labor mobilization during the reform period, thus seemingly violating the hypothesized inverse pattern. Collier and Collier explain these deviations based on crucial "flaws" in the strength of the Peruvian and Argentine oligarchies (pp. 748-49). Once these flaws are taken into account, the inverse relationship at work for the other six cases also makes sense for Peru and Argentina. In this way, Collier and Collier show how what initially appears to be a deviation in fact reflects the general inverse pattern once more appropriate measures are introduced.

Another work that relies extensively on ordinal comparison is Orloff's (1993) comparative-historical analysis of social provision for the elderly in Britain, Canada, and the United States. Orloff not only conceptualizes her own central explanatory variables as ordinal categories but also evaluates rival explanations using this type of assessment. For example, Orloff uses ordinal comparison to reject explanations of the relative timing of social provision that focus on changes in the number of aged persons present in society (pp. 47-48). Orloff argues that the expected relationship between age distribution and the timing of pension legislation is not supported by her cases. For example, Britain witnessed only a marginal increase in the number of elderly citizens but nevertheless enacted pensions at an early date. By contrast, Canada saw a very substantial increase in the elderly population but adopted old-age insurance at a relatively late date. For the United States, the elderly population exhibited intermediate to high levels, yet pensions were adopted at a very late date. Hence, there is no apparent relationship between the two variables, and Orloff eliminates size of the elderly population as a potential explanatory factor.

ORDINAL COMPARISON, NOMINAL COMPARISON, AND LARGE-N ANALYSIS

The logic of causation small-*N* researchers adopt when using ordinal comparison contrasts with the logic they adopt when using

nominal comparison. This contrast underlies much debate over the similarities and differences between small-*N* analysis and large-*N* analysis and has key implications for the question of whether methodological insights from large-*N* research are appropriate for small-*N* research.

In important respects, when using a *nominal* strategy of causal inference, small-*N* analysts employ an understanding of causation that is different from that commonly used in large-*N*, statistical research. Small-*N* researchers who use nominal comparison as a central strategy seek out patterns of causality in which the association between cause and effect is in certain respects fully predictable. By contrast, most contemporary large-*N* analysts assume the presence of random variation and believe that causation is partial and probabilistic (King et al. 1994:79-82). The two traditions thus often operate with different definitions of "causal effect." The probabilistic definition used in large-*N* analysis leads to a focus on the proportion of an outcome that can be attributed to a particular value on an explanatory variable.¹⁰ By contrast, the deterministic definition used with a nominal strategy in small-*N* analysis leads to a focus on the outcome generated by sufficient conditions and/or enabled by necessary conditions.

Because of these differences, small-*N* scholars who are committed to a deterministic approach to causation may be skeptical of efforts to use insights from statistical research as a basis for improving the quality of their research. Indeed, many of the key insights from statistical research about case selection and measurement are not fully relevant to small-*N* research if one is thinking about causality in terms of necessary and sufficient conditions. Thus, while King et al. (1994) seek to enhance qualitative research through insights from quantitative research, Ragin (1997, 1998), McKeown (1999), and Munck (1998) caution against using quantitative research as a basis for informing small-*N* research practices.

However, when using an *ordinal* strategy of causal inference, small-*N* analysts understand causality in a manner more compatible with large-*N* analysis: Causation is seen as probabilistic, and causal inference is understood as roughly the process whereby one estimates the proportion of an outcome that can be attributed to a particular value on an explanatory variable. Insights from statistical research

about case selection, measurement error, and the construction of causal theories are more applicable to small- N research when investigators employ ordinal comparison.

Given that some small- N researchers use both nominal and ordinal comparison, and even may use both techniques to evaluate the same causal pattern, it is important to ask whether it is meaningful and logically consistent to adopt two opposing understandings of causality in a single piece of research. There is a compelling argument that suggests this practice is not logically consistent. In particular, serious concerns may arise about the practice of first assessing a variable as a necessary or sufficient condition and then reassessing this variable at the same level of aggregation using ordinal comparison.

Necessary and sufficient causes do not typically follow a linear pattern when plotted against an outcome. Rather, the scatter plot for a variable that is a necessary or sufficient cause will be truncated along both the x and y axes; that is, there will be no cases with values above (or below) certain threshold points on the explanatory and outcome variables. For example, if a high value on some explanatory variable is necessary for a high value on some outcome, there will be no cases in the upper left portion of the scatter plot (i.e., there will be no cases with low values on the explanatory variable and high values on the outcome variable). By contrast, if a high value on some explanatory variable is sufficient for a high value on some outcome, there will be no cases in the lower right portion of the scatter plot (i.e., there will be no cases with high values on the explanatory variable and low values on the outcome variable). Thus, because in a two-by-two table there will routinely be cases in one of the off-diagonal cells, the pattern of causation underlying the idea of necessary and sufficient causation is not strictly a linear one.

Yet, when small- N researchers use ordinal analysis to assess hypotheses, they do generally test to see whether variable scores can be matched in a pattern that resembles linear change. For example, high values on an explanatory variable are typically evaluated to see if they are matched with high (or low) values on an outcome variable.¹¹ One would therefore not necessarily expect a necessary or sufficient condition to appear causally significant when evaluated through ordinal comparison. In short, it may be logically inconsistent to test

a given hypothesis using both nominal comparison and ordinal comparison.

This problem will not occur in small-*N* studies that use nominal and ordinal techniques to assess *different* causal relationships. Likewise, it is less of an issue for scholars who disaggregate variables when moving from nominal comparison to ordinal comparison. Rather, the problem primarily arises in studies that assess a given variable using both nominal and ordinal strategies without shifting levels of aggregation. In general, this problem can be avoided if analysts are explicit about whether they believe a causal pattern follows a linear model of change or some alternative model such as necessary and sufficient causation.

WITHIN-CASE ANALYSIS

In addition to nominal and ordinal cross-case comparison, small-*N* analysts rely extensively on within-case analysis. Within-case analysis examines multiple features of what was originally considered only a single case to assess whether associations developed through cross-case analysis are in fact causal. In making within-case evaluations, small-*N* analysts will often rely on nominal and ordinal measurement. However, whereas the nominal and ordinal strategies discussed above entail highly aggregated comparisons across cases, the procedures discussed in this section involve disaggregated comparisons within cases. Hence, a within-case strategy is distinguished from nominal and ordinal strategies in terms of level of aggregation.

In small-*N* research, within-case analysis is a tool specifically designed to compensate for limitations associated with cross-case methods. The most general type of within-case analysis is *pattern matching*, a procedure in which the analyst assesses cross-case associations in light of multiple within-case hypotheses. An important subtype of this procedure is *process tracing*, a technique in which the analyst attempts to locate the causal mechanisms linking a hypothesized explanatory variable to an outcome. Finally, a third technique—*causal narrative*—combines cross-case analysis and within-case analysis by comparing cases in terms of highly disaggregated sequences of processes and events that lead to outcomes.

PATTERN MATCHING

Causal patterns derived from cross-case comparison often suggest additional hypotheses about aspects of specific cases. Following a procedure that Campbell (1975) calls pattern matching, small-*N* analysts test these additional hypotheses, evaluating whether patterns derived from cross-case analysis can be matched with observations from within specific cases.¹² Campbell points out that pattern matching provides a powerful tool for theory falsification: Small-*N* researchers routinely find that their arguments cannot be sustained when within-case hypotheses are assessed (p. 182). Alternatively, if within-case observations are repeatedly consistent with a cross-case finding, researchers have stronger grounds for believing the cross-case finding is valid.

Small-*N* analysts pursue pattern matching using different levels of measurement. Both nominal measurement and ordinal measurement can be used in conjunction with pattern matching. Small-*N* researchers may also use interval measurement when assessing within-case hypotheses. Indeed, if a large number of within-case observations are measured at an interval level, small-*N* researchers may employ standard statistical methods with the pattern-matching procedure.

A nice example of the use of statistical research for the purpose of pattern matching is Goldstone's (1991) work on revolutions during the early-modern period. Goldstone's cross-case nominal argument suggests that demographic growth leads to revolutions by triggering structural crises (i.e., fiscal crises, elite-state and intraelite conflict, and mass opposition). To bolster this cross-case, small-*N* argument, Goldstone conceptualizes explanatory variables in terms of a large number of within-case quantitative measures and combines these measures into an overall "political stress indicator" that is evaluated statistically. These statistical evaluations are used as supplementary evidence to assess hypotheses that apply to a small number of cases, offering powerful confirmatory evidence in support of Goldstone's small-*N* argument.

In contrast to Goldstone's statistical analysis of within-case patterns, Luebbert (1991) uses ordinal and nominal comparison when employing pattern matching. For example, his argument that an alliance between the socialists and the middle peasantry (a "red-green"

alliance) caused social democracy in interwar Europe has multiple within-case implications. Some of these within-case implications entail ordinal propositions; for example, the governing social coalition will lack a high level of stability, the working class will exercise a high degree of autonomy from the state, and high levels of strikes and labor activism will develop (pp. 234-36). Other within-case implications entail nominal propositions; for example, socialists will not challenge the distribution of wealth in the countryside or try to mobilize the rural proletariat, and the middle peasantry will not provide a viable social base for the socialists (pp. 268-69, 272, 286-88). Luebbert's within-case analysis finds support for these hypotheses, significantly enhancing one's confidence that the cross-case argument is correct.

Regardless of the level of measurement employed (nominal, ordinal, interval), the additional leverage offered by pattern matching helps compensate for weaknesses of cross-case strategies. For example, one of the limitations of cross-case nominal methods is that several explanations may be supported by the data, leaving the analyst without a clear basis for deciding which explanatory factor is the most important. Pattern matching helps narrow the range of potential explanations by offering an additional means of eliminating variables. After variables are eliminated through pattern matching, small-*N* analysts are often left with much more parsimonious explanations.

Pattern matching is also a key tool for those analysts who seek to avoid the determinism of nominal methods. Analysts can use pattern matching to show that a relationship is causal despite the fact that a cross-case nominal comparison reveals one or more cases in which scores on the explanatory and outcome variables deviate from a general pattern of matching. For example, if only three out of four cases reflect a general pattern of matching on nominal explanatory and outcome variables, scholars who reject deterministic assumptions may nevertheless conclude that the pattern reflects causation if they find significant within-case support. Likewise, analysts may use pattern matching to argue that a relationship is not causal even though scores on an explanatory variable are perfectly matched with scores on an outcome variable in a cross-case nominal assessment. They can do so by showing how the cross-case pattern is not supported when assessed against multiple within-case patterns. Hence, small-*N* analysts who

choose to think about causation in probabilistic terms may use pattern matching as a basis for retaining explanatory variables that do not withstand deterministic cross-case nominal tests while rejecting other explanatory variables even though they do withstand such tests.

Finally, pattern matching can be a valuable supplement to cross-case ordinal comparison. Pattern matching can help analysts make a better judgment about the causal status of a relationship that is ambiguous when evaluated through cross-case ordinal analysis (as in Table 4). Likewise, pattern matching can call into question the findings of ordinal comparison, showing how an apparently causal relationship is in fact not causal when viewed in light of multiple within-case implications.

PROCESS TRACING

An important part of causal inference involves specifying the causal effect of a hypothesized explanatory variable. Yet, for many small-*N* analysts, an equally important part involves identifying the **causal mechanisms** that link an explanatory variable with an outcome variable (Blalock 1961:9; Elster 1989:4-7; Hedstrom and Swedberg 1998; Little 1991:15-19; Salmon 1984, chap. 5; Goldthorpe 1998). Causal mechanisms can be defined as the processes and intervening variables through which an explanatory variable exerts a causal effect on an outcome variable (Bennett 1997). Following George and McKeown (1985), the effort to **infer causality through the identification of causal mechanisms can be called process tracing.**¹³

Process tracing is often used to help the small-*N* analyst avoid mistaking a spurious correlation for a causal association. The problem of spuriousness arises when two correlated variables appear to be causally related but in fact are the product of an antecedent variable. In small-*N* research, cross-case comparative methods are often vulnerable to this problem. For example, when three temporally ordered variables are correlated in a sequence, small-*N* analysts have difficulty using cross-case methods to determine whether the sequence represents a causal path or a spurious correlation. The first variable in such a sequence is often perfectly correlated with both the second variable and the third variable. Small-*N* cross-case methods do not provide a strong basis for judging whether this first variable represents an

antecedent cause that explains away the presumed causal relationship between the second and third variables, or whether the first and third variables are correlated because of the presence of the second variable, in which case the idea of causal path makes sense. Process tracing can help the small-*N* analyst distinguish between these two possibilities by showing whether causal mechanisms link the variables together. Thus, if hypothesized causal mechanisms can be identified between the second and third variable through process tracing, the analyst has a basis for believing the sequence is a causal path; that is, the second variable has a real causal effect on the third variable. Alternatively, if causal mechanisms cannot be identified between the second and third variables, the analyst has grounds for believing the sequence may be a spurious correlation; that is, the second and third variables are correlated only because of the presence of the first antecedent variable.

Small-*N* analysts frequently argue that a correlation identified through cross-case analysis is not causal because mechanisms linking the presumed explanatory variable and outcome variable cannot be identified. For example, Skocpol's (1979:170-71) work on the origins of revolutions uses process tracing to reject causal variables—such as ideologically motivated vanguard movements—that were not eliminated through cross-case methods. Although ideologically motivated vanguard movements were present in her three cases of social revolution, Skocpol argues that they did not exert an important causal effect in bringing about revolutions. In particular, in contrast to what other scholars have hypothesized, Skocpol argues that vanguard movements were not responsible for triggering widespread revolts against landlords and state agents. Rather, vanguard movements were marginal to the central political processes that defined social revolutions, emerging on the scene only very late to take advantage of situations they did not create. Hence, Skocpol concludes that these movements were not a crucial cause of social revolutions in France, Russia, and China.

Likewise, Luebbert (1991) uses process tracing to eliminate the Moore-Gerschenkron thesis, which holds that fascist regimes result from the presence of a labor-repressive landed elite that is able to draw substantial lower-class rural support for fascism (pp. 308-9). Although there is a correlation between the presence/absence of a

repressive landed elite and the presence/absence of fascism, Luebbert suggests that the mechanisms through which this specific factor supposedly produces fascism are not supported by the historical record of the fascist cases. Thus, rural support for fascism was generally not present in areas where a landed elite predominated. Likewise, the evidence shows that the landed elites who could deliver large numbers of votes did not usually support fascism (pp. 308-9). In short, despite the correlation, Luebbert rejects the Moore-Gerschenkron hypothesis because it is not validated by process tracing.

Other scholars use process tracing not to eliminate causal factors but to support their own explanations. For example, Collier and Collier (1991) identify mechanisms linking different types of labor incorporation periods with different types of party systems. In their analysis of Colombia and Uruguay, Collier and Collier systematically identify the processes and events through which the incorporation pattern of "electoral mobilization by a traditional party" led to the party system outcome of "electoral stability and social conflict." These processes included a period in which the party that oversaw incorporation briefly maintained power, the gradual emergence of conservative opposition, a period of intense political polarization, a military coup, and, finally, the creation of a party system marked by stable electoral politics and social conflict. Each of these events acts as a mechanism linking labor incorporation with a particular party system outcome. The ability of Collier and Collier to show how these and other processes connected explanatory and outcome variables is crucial to the success of their argument. Indeed, although any small-*N* study can benefit from process tracing, it is an especially important tool for studies such as Collier and Collier's in which explanatory and outcome variables are separated by long periods of time.

A causal mechanism approach has received considerable attention and debate in the philosophy of science, especially among "scientific realists" (e.g., Harre 1972; Bhaskar 1975; Salmon 1984; McMullin 1984). Scholars in the social sciences responding to this discussion have themselves entered the debate, focusing on issues such as whether causal mechanisms are observable entities and whether causal mechanisms should be identified through general theories, statistical research, or qualitative comparative analysis (Kiser and Hechter 1991; King et al. 1994; Kiser 1996; Bennett 1997; Steinmetz

1998; Goldthorpe 1997, 1998; Somers 1998). In light of actual small-*N* research practice, however, these debates appear less as mutually exclusive positions and more as alternative approaches to the analysis of causal mechanisms. Thus, small-*N* researchers have understood causal mechanisms as both observable and unobservable entities and have used different techniques for determining whether causal mechanisms are in operation.

CAUSAL NARRATIVE

A final procedure illustrates how small-*N* researchers use cross-case comparisons of within-case chronologies as a basis for making causal inferences. With this technique of "causal narrative," to use Sewell's (1996) terminology, the analyst attempts to validate aggregated cross-case associations by separating variables into constituent sequences of disaggregated events and comparing these disaggregated sequences across cases. The purpose of unpacking aggregated variables through narrative is not simply to provide a contextualized description of cases; rather, the goal is to support a cross-case argument at a more disaggregated level.

This technique relies on historical narrative, which has received significant attention in recent methodological discussions (e.g., Abbott 1990, 1992; Aminzade 1992; Franzosi 1998; Griffin 1992, 1993; Haydu 1998; Isaac 1997; Mahoney 1999; Somers 1992; Stryker 1996). However, the procedures through which analysts decide whether a narrative account lends support to a cross-case causal pattern have not been well specified. Griffin's (1993; see also Heise 1989; Griffin and Korstad 1998) discussion of event-structure analysis is the most developed statement on how narrative can be wedded to causal inference. Event-structure analysis provides a formal apparatus for unpacking events and reconstituting their constituent parts as a causal interpretation of historical processes. This procedure can underpin causal narrative by identifying the causally linked processes that constitute highly aggregated variables in cross-case analysis (Mahoney 1999:1165-67). With causal narrative, the analyst compares event sequences across cases to determine whether cases can reasonably be seen as following aggregated causal patterns at a more fine-grained level. In this sense, causal narrative entails the matching of event

structures across cases (see Griffin and Ragin 1994:14-15; Sewell 1996:262).¹⁴ In addition, causal narrative can be used to show how two or more cases that are marked by important differences in causal processes at an aggregated level of analysis are also characterized by substantially different event structures at a disaggregated level.

A good example of the use of causal narrative to compare event structures is found in Skocpol's (1979) work on social revolutions. Many of Skocpol's key explanatory variables are actually made up of numerous causally linked processes. Likewise, the outcome of social revolution is itself composed of a series of causally connected events. These constituent processes represent an event-structure pattern that could be formally diagrammed and compared across cases (see Mahoney 1999). Although Skocpol does not carry out a formal mapping of event structures, she does implicitly compare the event structure of her cases to judge whether they follow a similar causal logic at a disaggregated level. According to Sewell (1996), Skocpol's ability to show a similar event sequence is at work in each case of social revolution greatly contributes to the persuasiveness of her argument.

An interesting example of a work that uses causal narrative to contrast event-structure sequences is Yashar's (1997) excellent analysis of the origins of democracy in Costa Rica in 1949 and authoritarianism in Guatemala in 1954. Yashar notes that both Costa Rica and Guatemala experienced major democratic and social reform periods in the 1940s and 1950s. However, her narrative shows that these reform periods were actually composed of quite different event processes, and these different processes were highly consequential for the development of contrasting regimes. Thus, Yashar's narrative shows how particular actor choices about coalitional allies triggered differing reform efforts, reactions, and counterreactions. These differences in event sequences lend support to Yashar's overarching argument, which stresses the importance of political coalitions and alliance patterns.

Most small-*N* analysts use causal narrative as an informal technique presented through "stories" of event processes. It remains to be seen whether causal narrative can be more formally employed through event-structure diagrams that explicitly map each step and logical connection in a narrative argument. Griffin's (1993) work on event-structure analysis suggests that formally diagramming narratives can be complicated when a large number of events are considered. Yet,

without such formal diagramming, the procedures through which analysts compare and contrast event-structure sequences across cases cannot be easily evaluated. The complex trade-offs between the rhetorical clarity of informal narrative presentations and the rigor of explicitly diagrammed narrative accounts warrant further attention from methodologists.

CONCLUSION

This article has discussed the logic underlying three basic strategies for causal inference in small-*N* analysis: nominal comparison, ordinal comparison, and within-case analysis. These strategies do not necessarily represent the full spectrum of approaches through which small-*N* analysts evaluate causal hypotheses. However, they do stand out as major identifiable tools used in this research tradition. By way of conclusion, I offer three suggestions for small-*N* analysts about how to improve their use of these strategies. In addition, I offer two suggestions for scholars who seek to use the leverage of statistical methods as a basis for evaluating the research techniques of small-*N* analysts.

First, small-*N* analysts need to be more explicit about whether they are using nominal comparison as an informal strategy simply to develop an initial idea of whether variables are roughly matched or as a central strategy of hypothesis evaluation in which potential explanations are systematically eliminated. If analysts are using nominal comparison as a central strategy of inference, they must recognize that these methods eliminate potential explanations through a deterministic understanding of causation built around the ideas of necessary and sufficient causation. Small-*N* analysts who reject the notions of necessary and sufficient causation should avoid using nominal methods as a central basis for eliminating rival hypotheses.

Second, small-*N* analysts need to be more explicit about the ways in which they use ordinal comparison. Although I believe ordinal comparison is widely employed in small-*N* research, few researchers are up front about their use of this procedure. This lack of explicitness has prompted a misleading portrait of small-*N* research as entailing the study of differences of “kind” rather than differences of “degree.”

Furthermore, scholars who use both nominal and ordinal strategies of cross-case analysis need to do a better job of highlighting the relationship between the two strategies. Is ordinal analysis used to supplement a primarily nominal assessment, or is ordinal analysis the principal strategy of causal inference? Is ordinal analysis used at the same level of aggregation as nominal analysis, or are variables disaggregated in the move from nominal analysis to ordinal analysis? Small-*N* analysts would do well to explicitly address questions such as these.

Third, small-*N* researchers need to acknowledge the specific type or types of within-case analysis they use. There is no inherent reason why within-case analysis must remain an underspecified and implicit strategy. Indeed, researchers should strive to address the key issues that typically arise with the application of particular types of within-case analysis. For example, when using pattern matching, small-*N* analysts should be more self-conscious about the procedure of causal inference (i.e., nominal, ordinal, or statistical) they follow at the within-case level. Likewise, when employing process tracing to locate causal mechanisms, scholars should offer as precise an account as possible of the processes that are hypothesized to link an explanatory variable with an outcome variable. The testable implications of these processes should also be stated as clearly as possible. Finally, with regard to causal narrative, more methodological analysis needs to be brought to bear on the actual ways in which event structures are organized and on their relationship to arguments developed through cross-case comparative methods.

I would like to close with two suggestions for scholars who seek to evaluate small-*N* research by drawing on insights from large-*N* research. First, it is important that large-*N* scholars assess small-*N* research in light of the interplay of the multiple methods that are used in this research tradition. Despite a prominent place in the literature, efforts aimed at criticizing small-*N* analysis from a large-*N* perspective that focus exclusively on one strategy of causal inference have been misleading. Future critiques and commentaries will be much more useful if they focus on the full range of procedures employed in small-*N* research.

Second, large-*N* researchers who seek to offer methodological guidance to small-*N* researchers must recognize that insights from quantitative research cannot always be directly translated into helpful

advice for small-*N* researchers. For example, insights from linear regression analysis are often counterproductive when applied to small-*N* studies that employ nominal comparison. Likewise, it remains to be seen how statistical research can best inform small-*N* efforts to identify causal mechanisms. The real methodological task of unifying quantitative and qualitative methods involves not merely prodding small-*N* researchers to learn more about the procedures used in large-*N* research. Rather, it also entails encouraging large-*N* researchers to learn more about the diverse strategies of causal inference used in small-*N* research.

NOTES

1. For criticisms of Mill's method of agreement and method of difference, see Goldthorpe (1997), Lieberman (1991, 1994), and Nichols (1986). In defense of these methods, see Savolainen (1994) and Skocpol (1984, 1986). See also Ragin (1987).

2. With the method of agreement, the analyst attempts to establish that cases that share a common outcome also share common hypothesized causal factors, despite varying in other significant ways.

3. The analyst using method of difference contrasts cases in which an outcome under investigation and hypothesized causal factors are present to other cases in which both the outcome and the hypothesized causes are absent, even though the cases may be similar in many other respects.

4. The Boolean minimization procedure through which causal expressions are reduced eliminates potential *necessary* causes among cases that share the same outcome. Final combinations of causal factors are understood to be *sufficient* for the occurrence of the outcome. In addition, by looking at final combinations of explanatory variables, it is possible to state whether any single factor is a necessary or sufficient cause.

5. The degrees-of-freedom problem is probably the most common criticism of small-*N* analysis. For different perspectives, see Campbell (1975), Collier (1993), Goldthorpe (1997), King et al. (1994), Lieberman (1991), Lijphart (1971, 1975), and Nichols (1986).

6. The method of concomitant variation can be used deterministically. In this case, the analyst infers causation only when a pattern of perfect ordinal matching is present. The underlying assumption is that a given ordinal score on the explanatory variable is a necessary and sufficient condition for a given ordinal score on the outcome variable.

7. This assumes that cases are selected based on their score on the outcome variable. If cases are selected on the outcome variable, then the probability of finding covariation between cause and outcome depends on the distribution of scores on the causal variable in the full population of cases.

8. This percentage was calculated using simple probability theory and the assumption that cause X is present half of the time in the full population of cases.

9. The calculation of this percentage takes into consideration the degree to which cause and outcome do not covary (e.g., given the overall pattern in Table 3, a case with a score of 1 on the explanatory variable and a score of 4 on the outcome variable is a bigger "miss" than a case with a score of 2 on the explanatory variable and 3 on the outcome variable). The probability calcula-

tion again assumes that cases were selected based on their score on the outcome variable. The Spearman rank-order coefficient for Table 3 is .955.

10. See King et al. (1994:76-82), who draw extensively on Holland (1986).

11. Strictly speaking, ordinal analysis does not entail the conceptualization of continuously measured variables and, thus, cannot be used to assess whether variables follow a linear pattern.

12. See also Collier (1993), Eckstein (1975), and George (1979).

13. Because causal mechanisms are within-case implications of cross-case patterns, process tracing is actually a particular application of pattern matching.

14. Causal narrative cannot be used to assess cases that arrive at a given outcome through different causal processes. When this is true, one would expect the cases to be characterized by different—not similar—event-structure sequences.

REFERENCES

- Abbott, Andrew. 1990. "Conceptions of Time and Events in Social Science Methods: Causal and Narrative Approaches." *Historical Methods* 23:140-50.
- . 1992. "From Causes to Events: Notes on Narrative Positivism." *Sociological Methods & Research* 20:428-55.
- Amenta, Edwin and Jane D. Poulsen. 1994. "Where to Begin: A Survey of Five Approaches to Selecting Independent Variables for Qualitative Comparative Analysis." *Sociological Methods & Research* 23:22-53.
- Aminzade, Ronald. 1992. "Historical Sociology and Time." *Sociological Methods & Research* 20:456-80.
- Bennett, Andrew. 1997. "Lost in the Translation: Big (N) Misinterpretations of Case Study Research." Presented at the 38th Annual Convention of the International Studies Association, March 18-22, Toronto.
- Bhaskar, Roy. 1975. *A Realist Theory of Science*. London: Leeds Books.
- Blalock, Hubert M. 1961. *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.
- Campbell, Donald T. 1975. "'Degrees of Freedom' and the Case Study." *Comparative Political Studies* 8:178-93.
- Collier, David. 1993. "The Comparative Method." Pp. 105-19 in *Political Science: The State of the Discipline II*, edited by Ada Finifter. Washington, DC: American Political Science Association.
- Collier, David and Robert Adcock. 1999. "Democracy and Dichotomies: A Pragmatic Approach to Choices About Concepts." *Annual Review of Political Science* 2:537-65.
- Collier, David and James Mahoney. 1996. "Insights and Pitfalls: Selection Bias in Qualitative Research." *World Politics* 49:56-91.
- Collier, Ruth Berins and David Collier. 1991. *Shaping the Political Arena: Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America*. Princeton, NJ: Princeton University Press.
- DeFelice, E. Gene. 1986. "Causal Inference and Comparative Methods." *Comparative Political Studies* 19:415-37.
- Dion, Douglas. 1998. "Evidence and Inference in the Comparative Case Study." *Comparative Politics* 30:127-46.
- Downing, Brian M. 1992. *The Military Revolution and Political Change: Origins of Democracy and Autocracy in Early Modern Europe*. Princeton, NJ: Princeton University Press.

- Eckstein, Harry. 1975. "Case Studies and Theory in Political Science." Pp. 79-138 in *Handbook of Political Science* 7, edited by Fred Greenstein and Nelson W. Polsby. Reading, MA: Addison-Wesley.
- Elster, Jon. 1989. *Nuts and Bolts for the Social Sciences*. Cambridge, UK: Cambridge University Press.
- Ertman, Thomas. 1997. *Birth of the Leviathan: Building States and Regimes in Medieval and Early Modern Europe*. Cambridge, UK: Cambridge University Press.
- Esping-Andersen, Gosta. 1990. *The Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.
- Franzosi, Roberto. 1998. "Narrative as Data: Linguistic and Statistical Tools for the Qualitative Study of Historical Events." *International Review of Social History* 43:81-104.
- Fulbrook, Mary. 1983. *Piety and Politics: Religion and the Rise of Absolutism in England, Wurttemberg and Prussia*. Cambridge, UK: Cambridge University Press.
- Geddes, Barbara. 1990. "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics." Pp. 131-50 in *Political Analysis*, vol. 2, edited by James A. Stimson. Ann Arbor: University of Michigan Press.
- George, Alexander L. 1979. "Case Studies and Theory Development: The Method of Structured, Focused Comparison." Pp. 43-68 in *Diplomacy: New Approaches in History, Theory, and Policy*, edited by P. G. Lauren. New York: Free Press.
- George, Alexander L. and Timothy J. McKeown. 1985. "Case Studies and Theories of Organizational Decision Making." *Advances in Information Processing in Organizations* 2:21-58.
- Goldstone, Jack A. 1991. *Revolution and Rebellion in the Early Modern World*. Berkeley: University of California Press.
- . 1997. "Methodological Issues in Comparative Macrosociology." *Comparative Social Research* 16:107-20.
- Goldthorpe, John H. 1997. "Current Issues in Comparative Macrosociology: A Debate on Methodological Issues." *Comparative Social Research* 16:1-26.
- . 1998. "Causation, Statistics and Sociology." Unpublished manuscript, Nuffield College, Oxford University.
- Goodwin, Jeff. 2000. *No Other Way Out: States and Revolutionary Movements, 1945-1991*. Cambridge, UK: Cambridge University Press.
- Goodwin, Jeff and Theda Skocpol. 1989. "Explaining Revolutions in the Contemporary Third World." *Politics and Society* 17:489-507.
- Gould, Andrew C. 1999. *The Origins of Liberal Dominance: State, Church, and Party in Nineteenth Century Europe*. Ann Arbor: University of Michigan Press.
- Griffin, Larry J. 1992. "Temporality, Events, and Explanation in Historical Sociology: An Introduction." *Sociological Methods & Research* 20:403-27.
- . 1993. "Narrative, Event-Structure, and Causal Interpretation in Historical Sociology." *American Journal of Sociology* 98:1094-133.
- Griffin, Larry J. and Robert R. Korstad. 1998. "Historical Inference and Event-Structure Analysis." *International Review of Social History* 43:145-65.
- Griffin, Larry and Charles C. Ragin. 1994. "Some Observations on Formal Methods of Qualitative Analysis." *Sociological Methods & Research* 23:4-21.
- Harre, R. 1972. *The Philosophies of Science*. London: Oxford University Press.
- Haydu, Jeffrey. 1998. "Making Use of the Past: Time Periods as Cases to Compare and as Sequences of Problem Solving." *American Journal of Sociology* 104:339-71.
- Hedstrom, Peter and Richard Swedberg, eds. 1998. *Social Mechanisms: An Analytical Approach to Social Theory*. New York: Cambridge University Press.

- Heise, David. 1989. "Modeling Event Structures." *Journal of Mathematical Sociology* 14:139-69.
- Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81:945-60.
- Isaac, Larry W. 1997. "Transforming Localities: Reflections on Time, Causality, and Narrative in Contemporary Historical Sociology." *Historical Methods* 30:4-12.
- Katznelson, Ira. 1997. "Structure and Configuration in Comparative Politics." Pp. 81-112 in *Comparative Politics: Rationality, Culture, and Structure*, edited by Mark Irving Lichbach and Alan S. Zuckerman. Cambridge, UK: Cambridge University Press.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.
- Kiser, Edgar. 1996. "The Revival of Narrative in Historical Sociology: What Rational Choice Can Contribute." *Politics and Society* 24:249-71.
- Kiser, Edgar and Michael Hechter. 1991. "The Role of General Theory in Comparative-Historical Sociology." *American Journal of Sociology* 97:1-30.
- Lieberson, Stanley. 1991. "Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases." *Social Forces* 70:307-20.
- . 1994. "More on the Uneasy Case for Using Mill-Type Methods in Small-N Comparative Studies." *Social Forces* 72:1225-37.
- . 1998. "Causal Analysis and Comparative Research: What Can We Learn From Studies Based on a Small Number of Cases." Pp. 129-45 in *Rational Choice Theory and Large-Scale Data Analysis*, edited by Hans-Peter Blossfeld and Gerald Prein. Boulder, CO: Westview.
- Lijphart, Arend. 1971. "Comparative Politics and the Comparative Method." *American Political Science Review* 65:682-93.
- . 1975. "The Comparable Cases Strategy in Comparative Research." *Comparative Political Studies* 8:158-77.
- Little, Daniel. 1991. *Varieties of Social Explanation: An Introduction to the Philosophy of Social Science*. Boulder, CO: Westview.
- Luebbert, Gregory M. 1987. "Social Foundations of Political Order in Interwar Europe." *World Politics* 39:449-78.
- . 1991. *Liberalism, Fascism, or Social Democracy: Social Classes and the Political Origins of Regimes in Interwar Europe*. New York: Oxford University Press.
- Mahoney, James. 1999. "Nominal, Ordinal, and Narrative Appraisal in Macrocausal Analysis." *American Journal of Sociology* 104:1154-96.
- McKeown, Timothy J. 1999. "Case Studies and the Statistical Worldview." *International Organization* 53:161-90.
- McMullin, Ernan. 1984. "Two Ideals of Explanation in Natural Science." *Midwest Studies in Philosophy* 9:205-20.
- Mill, John Stuart. [1843] 1974. *A System of Logic*. Toronto: University of Toronto Press.
- Munck, Gerardo L. 1998. "Canons of Research Design in Qualitative Analysis." *Studies in Comparative International Development* 33:18-45.
- Nichols, Elizabeth. 1986. "Skocpol on Revolution: Comparative Analysis vs. Historical Con-juncture." *Comparative Social Research* 9:163-86.
- Orloff, Ann Shola. 1993. *The Politics of Pensions: A Comparative Analysis of Britain, Canada, and the United States, 1880-1940*. Madison: University of Wisconsin Press.
- Przeworski, Adam and Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley.
- Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

- . 1997. "Turning the Tables: How Case-Oriented Research Challenges Variable-Oriented Research." *Comparative Social Research* 16:27-42.
- . 1998. "Case-Oriented Research and the Study of Social Action." Pp. 158-68 in *Rational Choice Theory and Large-Scale Data Analysis*, edited by Hans-Peter Blossfeld and Gerald Prein. Boulder, CO: Westview.
- Ragin, Charles and David Zaret. 1983. "Theory and Method in Comparative Research: Two Strategies." *Social Forces* 61:731-54.
- Rueschemeyer, Dietrich and John D. Stephens. 1997. "Comparing Historical Sequences—A Powerful Tool for Causal Analysis." *Comparative Social Research* 17:55-72.
- Salmon, Wesley. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.
- Sartori, Giovanni. 1987. *Theory of Democracy Revisited*. Chatham, NJ: Chatham House.
- Savolainen, Jukka. 1994. "The Rationality of Drawing Big Conclusions Based on Small Samples: In Defense of Mill's Methods." *Social Forces* 72:1217-24.
- Sewell, William H., Jr. 1996. "Three Temporalities: Toward an Eventful Sociology." Pp. 245-80 in *The Historic Turn in the Human Sciences*, edited by Terrence J. McDonald. Ann Arbor: University of Michigan Press.
- Skocpol, Theda. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge, UK: Cambridge University Press.
- . 1984. "Emerging Agendas and Recurrent Strategies in Historical Sociology." Pp. 356-91 in *Vision and Method in Historical Sociology*, edited by Theda Skocpol. Cambridge, UK: Cambridge University Press.
- . 1986. "Analyzing Causal Configurations in History: A Rejoinder to Nichols." *Comparative Social Research* 9:187-94.
- . 1992. *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*. Cambridge, MA: Harvard University Press.
- Skocpol, Theda and Margaret Somers. 1980. "The Uses of Comparative History in Macrosocial Inquiry." *Comparative Studies in Society and History* 22:174-97.
- Smelser, Neil. 1976. *Comparative Methods in the Social Sciences*. Englewood Cliffs, NJ: Prentice Hall.
- Sobel, Michael E. 1995. "Causal Inference in the Social and Behavioral Sciences." Pp. 1-38 in *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, edited by Gerhard Arminger, Clifford C. Clogg, and Michael E. Sobel. New York: Plenum.
- Somers, Margaret R. 1992. "Narrativity, Narrative Identity, and Social Action: Rethinking English Working-Class Formation." *Social Science History* 16:591-630.
- . 1998. "'We're No Angels': Realism, Rational Choice, and Relationality in Social Science." *American Journal of Sociology* 104:722-84.
- Steinmetz, George. 1998. "Critical Realism and Historical Sociology." *Comparative Studies in Society and History* 40:170-86.
- Stryker, Robin. 1996. "Beyond History Versus Theory: Strategic Narrative and Sociological Explanation." *Sociological Methods & Research* 24:304-52.
- Wickham-Crowley, Timothy. 1992. *Guerrillas and Revolution in Latin America: A Comparative Study of Insurgents and Regimes Since 1956*. Princeton, NJ: Princeton University Press.
- Yashar, Deborah J. 1997. *Demanding Democracy: Reform and Reaction in Costa Rica and Guatemala, 1870s-1950s*. Stanford, CA: Stanford University Press.

James Mahoney is an assistant professor of sociology at Brown University. His interests are Latin American politics, historical sociology, and qualitative methodology. His dissertation was cowinner of the 1998 Gabriel Almond Prize of the American Political Science Association.