Qualitative Methodology and Comparative Politics
James Mahoney
Comparative Political Studies 2007; 40; 122
DOI: 10.1177/0010414006296345

The online version of this article can be found at:
http://cps.sagepub.com/cgi/content/abstract/40/2/122

Published by:
SAGE
http://www.sagepublications.com

Additional services and information for Comparative Political Studies can be found at:

Email Alerts: http://cps.sagepub.com/cgi/alerts

Subscriptions: http://cps.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://cps.sagepub.com/cgi/content/refs/40/2/122
Leading methods for pursuing qualitative research in the field of comparative politics are discussed. On one hand, qualitative researchers in this field use a variety of methods of theory development: procedures for generating new hypotheses, tools for pursuing conceptual innovation, and techniques for identifying populations of homogeneous cases. On the other hand, they employ both within-case and cross-case methods of theory testing. Within-case methods include techniques for identifying intervening mechanisms and testing multiple observable implications of theories. Cross-case methods include a host of approaches for assessing hypotheses about necessary and sufficient causes. The article discusses the distinctive leverage offered by qualitative research for addressing questions in comparative politics.

Keywords: case selection; theory generation; theory testing; causal inference

Beginning in the 1990s, the field of comparative politics saw an unprecedented wave of publications concerning qualitative and small-N methods (Bennett & Elman, 2006a; Munck, 1998). These publications built on work on comparative methodology dating to the 1970s (e.g., Lipjhart, 1971, 1975; Przeworski & Teune, 1970; Smelser, 1973, 1976). However, whereas the earlier work often viewed qualitative methodology as advancing a set of “last-resort” techniques that should be employed only when other methods (e.g., statistical methods) are not appropriate, the current work emphasizes the distinctive advantages of qualitative research. This new emphasis corresponds to research practices in the field. Students of comparative politics frequently turn to qualitative methods instead of or in combination with alternative techniques because they believe that qualitative methods are essential for addressing many substantive questions of interest.

Author's Note: An earlier version of this article was presented at the 2005 Annual Meetings of the American Political Science Association, Washington, D.C., September 1 to 4. For helpful comments, I thank James Caporaso, Richard Snyder, Larkin Terrie, and David Waldner. This material is based on work supported by the National Science Foundation under Grant 0093754.
Today, scholars using qualitative methods explore all of the major substantive topics in comparative politics. For evidence, one can point to influential work across the any of the key areas of the field: democracy and authoritarianism (e.g., Collier, 1999; Haggard & Kaufman, 1995; Huntington, 1991; Linz & Stepan, 1996; Rueschemeyer, Stephens, & Stephens, 1992), economic growth (e.g., Amsden, 2003; Evans, 1995; Kohli, 2004; Wade, 1990; Waterbury, 1993), market-oriented reform and regulation (e.g., Ekiert & Hanson, 2003; Haggard & Kaufman, 1992; Hall & Soskice, 2001; Kitschelt, 1994; Nelson, 1990; Vogel, 1996; Weyland, 2002), state building (e.g., Downing, 1992; Ertman, 1997; Tilly, 1990; Waldner, 1999), nationalism and ethnicity (e.g., Brubaker, 1992; Haas, 1997; Lustick, 1993; Marx, 1998; Varshney, 2002; Yashar, 2005), violence and state collapse (e.g., Boone, 2003; Reno, 1998), social revolutionary change (e.g., Colburn, 1994; Goodwin, 2001; Parsa, 2000), social movements (e.g., Goldstone, 2003; McAdam, Tarrow, & Tilly, 2001; Tarrow, 1994), electoral and party systems (e.g., Collier & Collier, 1991; Kitschelt, Masfeldova, Markowski, & Tóka, 1999; Mainwaring & Scully, 1995), and social policy (e.g., Esping-Andersen, 1990; Hacker, 2002; Hicks, 1999; Immergut, 1992; Pierson, 1994; Skocpol, 1992). Although this listing only scratches the surface of huge literatures, it does suggest the prominent place for research that uses qualitative methods.

Studies that are associated with alternative methodologies also suggest the centrality of qualitative methodology. For example, rational choice works often draw on qualitative methods to help formulate assumptions and test hypotheses. This is especially true in the field of comparative politics, where analysts usually employ rational actor assumptions in relatively nonmathematical ways to generate hypotheses that apply to a small number of cases (e.g., Bates, 1981; Geddes, 2003; Laitin, 1999; see also Munck, 2001). Indeed, the recent project of “analytic narratives” is an explicit attempt to meld the virtues of qualitative research with the assumptions of rational choice theory (Bates, Greif, Levi, Rosenthal, & Weingast, 1998). Likewise, statistical work in comparative politics increasingly uses qualitative case studies as a supplementary mode of causal inference (e.g., Huber & Stephens, 2001; Kalyvas, 2006; Lieberman, 2003; Posner, 2005; Swank, 2002; Wilkinson, 2004; see also Lieberman, 2005). In fact, in some of these multimethod studies, the qualitative case studies provide the main leverage for causal inference.

The increased focus on qualitative methodology has been accompanied by excellent synthetic articles that discuss different strands and various proposals related to the comparative method (e.g., Bennett & Elman, 2006b; Collier, 1993; Munck, 1998; Ragin, Berg-Schlosser, & de Muer, 1996). Yet the publication of no less than 10 major books on qualitative methodology in the past 5 years has made it difficult for scholars of comparative politics to...
keep apace with rapid developments in this field (Abbott, 2001; Brady & Collier, 2004; Geddes, 2003; George & Bennett, 2005; Gerring, 2001, 2006; Goertz, 2006; Mahoney & Rueschemeyer, 2003; Pierson, 2004; Ragin, 2000). This essay brings together recent insights and debates about qualitative methodology in comparative politics through an examination of two broad areas: theory development and theory testing.

Theory Development

In other subfields of political science, researchers may commonly work within general research programs that provide base-level assumptions for formulating testable theories. For example, scholarship in the field of international relations is often organized around research programs such as neo-realism, liberalism, and constructivism (Bennett & Elman, 2007 [this issue]; see also Elman & Elman, 2003). But in comparative politics, analysts usually do not draw on such encompassing research programs. Instead, they find theoretical inspiration in a wide variety of orienting approaches—strategic choice models, state-centric approaches, patron-client models, theories of international dependency, and many more—that emphasize certain key causal factors but that lack the generality that we normally associate with a research program or theoretical paradigm. Comparative analysts—whether rational choice theorists, structuralists, or culturalists (Lichbach & Zuckerman, 1997)—usually combine different elements of diverse research orientations to formulate their theories.

Working outside of singular and overarching theoretical programs, comparative analysts are especially likely to begin their research without well-developed and readily testable hypotheses. Methodological tools for developing testable hypotheses therefore are of great utility to comparativists who seek to move from research topics to specific propositions. Quantitative methodology is relatively silent about the procedures analysts might follow to formulate the sets of hypotheses that constitute testable theories, though some specification searches do allow for exploratory research of this nature. For its part, qualitative methodology offers its own various tools for framing research questions and formulating testable hypotheses.

Novel Hypothesis Formulation

Perhaps no one disagrees that over the years qualitative researchers have formulated interesting hypotheses in the field of comparative politics. We only need to think of our most provocative and recognizable propositions,
such as Michels’s (1911/1999) iron law asserting that “who says organization says oligarchy.” Moore’s (1966) claim of “no bourgeois, no democracy,” and Skocpol’s (1979) structural argument that “revolutions are not made, they come.” However, what is often less recognized is the role of qualitative data collection—interviews, participant observation, immersion in secondary sources, and archival research—in the development of novel hypotheses. Consider, for example, Scott’s (1985) famous book *Weapons of the Weak*. Scott developed his critique of the Gramscian hypothesis of false consciousness and formulated an alternative theory of everyday forms of peasant resistance above all through the close inspection of and engagement with a particular village (designated as Sedaka) in Malaysia during 2 years of field research. As Scott makes abundantly clear (pp. xvii-xix), the major conceptual innovations of his book grew out of observations and participation in this village. Indeed, without case-intensive knowledge of actual villagers, it would have been nearly impossible for him to discover the specific ways in which nonrebellious peasants recognize their “objective” interests and act on them accordingly. The book thus stands as an excellent example of what D. Collier (1999) calls “extracting new ideas at close range.”

Writings on qualitative methodology—both old and new—have taken steps toward formally codifying the procedures involved in generating new hypotheses. Perhaps the most widely discussed technique is the use of a “deviant case study” design to formulate new hypotheses (Eckstein, 1975; Emigh, 1997; George & Bennett, 2005; Lijphart, 1971). Deviant cases are observations with outcomes that do not conform to theoretical predictions. Qualitative analysts frequently study these cases intensely to understand the reasons why they defy theory, often in the process discovering novel hypotheses that can be tested more generally. Qualitative deviant case studies in comparative politics include very well-known and theoretically innovative works, such as Lipset, Trow, and Coleman’s (1956) *Union Democracy*, Lijphart’s (1968) *The Politics of Accommodation*, Johnson’s (1982) *MITI and the Japanese Miracle*, and Skocpol’s (1992) *Protecting Soldiers and Mothers*. Although deviant case analysis almost always involves close qualitative inspection, it can be creatively combined with quantitative research. In particular, regression analysis can be used to identify outlier cases that are then appraised with qualitative methods (for discussions and examples, see Coppelde, 2001; Lieberman, 2005).

A second technique of theory generation involves the use of comparisons to formulate new hypotheses. Qualitative analysts juxtapose multiple features of cases with one another, including features that they discover in the course of their analysis (Campbell, 1975; Ragin, 1987). Indeed, almost inevitably as a byproduct of contextualized comparisons, new concepts and explanatory
hypotheses are developed in qualitative analysis. By contrast, although statistical approaches certainly also systematically compare aspects of cases, they lack an obvious means of generating new hypotheses through the discovery of new variables. As George and Bennett (2005) put it, “Unless statistical researchers do their own archival work, interviews, or face-to-face surveys with open-ended questions in order to measure the values of the variables in their model, they have no unproblematic inductive means of identifying left-out variables” (p. 21). That is, unless statistical researchers turn to qualitative research of some kind, they are likely to include only those variables present from the start of the research.

Finally, qualitative research is a leading site of hypothesis creation in comparative politics because it facilitates the study of over-time data and a concern with temporal processes. Especially in the field of comparative-historical analysis, researchers take seriously the unfolding of events over time, and they often formulate new hypotheses that stress how the temporal intersection or duration of variables is decisive for outcomes (Mahoney & Rueschemeyer, 2003; see also Abbott, 2001; Pierson, 2004). It is thus no accident that qualitative researchers have been the ones who have developed techniques to analyze critical junctures and path-dependent processes of change (Bennett & Elman, 2006a; Collier & Collier, 1991; Lipset & Rokkan, 1968; Mahoney, 2000; Pierson, 2004; Thelen, 2003). These techniques, in turn, have inspired some of the most interesting theories of long-run political change in the field, such as Collier and Collier’s (1991) hypotheses about the effects of labor incorporation periods on long-run electoral dynamics in Latin America, Marx’s (1998) contention that the Civil War and Boer War in the United States and South Africa triggered coalitional patterns that forged reactionary racial orders for decades to come, and Thelen’s (2004) argument that the long-run evolution of the German Handicraft Protection Law of 1897, originally designed to shore support among a reactionary artisanal class, was critical to the development of contemporary Germany’s vocational training system.

In short, although theory development is partly driven by creativity and the scholarly imagination (see Munck & Snyder, 2007), it is also grounded in certain methodological techniques commonly practiced in qualitative research, including the close inspection of deviant cases, the use of highly contextualized comparisons, and the study of temporal processes and long-run analysis.

**Concepts and Measurement**

Concepts are central to theoretical innovation. The discovery of a new explanation or the development of a new research puzzle often goes hand
in hand with conceptual formulation and reformulation. In the field of comparative politics, qualitative researchers probably have been the scholars most persistently concerned with the definition (as opposed to quantitative operationalization) of the field’s key concepts. We see this, for example, by considering the range of articles by qualitative analysts on particular concepts: corporatism (D. Collier, 1999), democracy (D. Collier & Levitsky, 1997; Schmitter & Karl, 1991), ideology (Gerring, 1997), institutionalization (Levitsky, 1998), peasant (Kurtz, 2000), political culture (Gerring, 2003), revolution (Kotowski, 1984), and structure (Sewell, 1992). Moreover, through typological analysis (Elman, 2005; George & Bennett, 2005), qualitative researchers routinely develop conceptual distinctions that play a major role in driving theories of comparative politics. These typologies include, to name only a few, Linz and Stepan’s (1996) specification of types of regimes (e.g., democratic, authoritarian, totalitarian), Esping-Andersen’s (1990) types of welfare states (e.g., Christian, liberal, social-democratic), and Evan’s (1995) typology of states (e.g., predatory, developmental, intermediary).

Methodological writings have uncovered some of the reasons why there is an affinity between qualitative research and conceptual development. One answer is that qualitative researchers match background understandings of concepts with fine-grained evidence from cases (Adcock & Collier, 2001; Ragin, 1987). This process of matching, which often proceeds through many rounds of iteration, stimulates new conceptual understandings. For example, O’Donnell’s (1973, 1979, 1994) creation of various new concepts over the years—bureaucratic-authoritarianism, delegative democracy, horizontal accountability, and brown areas—has always occurred in conjunction with the close qualitative examination of actual cases. To capture a changing political reality in Latin America, O’Donnell has found it essential to invent new concepts—something quite different than devising new quantitative measures for existing concepts.

A growing methodological literature explicitly on concepts also has promoted conceptual innovation in qualitative research. The starting point of this literature is Sartori’s (1970, 1984) work, which explores concept formation through “checklist” definitions that treat conceptual attributes as individually necessary and jointly sufficient for conceptual membership. Sartori’s insights about an inverse relationship between a concept’s intension and extension provides the basis for further insights concerning “conceptual stretching” and the broader semantic fields of concepts. More recently, D. Collier and collaborators (Collier & Levitsky, 1997; Collier & Mahon, 1993) have drawn on ideas developed in cognitive science and linguistic philosophy to explore the family resemblance approach to concepts and tools for analyzing “radial” concepts (see also Lakoff, 1987). Additional approaches available to comparative
researchers include Gerring’s (2003) “min-max strategy” and Goertz’s (2006) major new work built around necessary and sufficient conditions.

The close examination of cases in qualitative research further helps comparative analysts avoid measurement error, a goal shared by scholars of all orientations in political science. At the most basic level, qualitative research is essential to good measurement because, as Brady (2004) notes, “qualitative comparisons are the basic building blocks of any approach to measurement” (p. 63). Moreover, measurement is enhanced when operational definitions and indicators can be refined in light of detailed case knowledge (Adcock & Collier, 2001). For example, in their famous comparison of famines and poverty in China and India, Dreze and Sen (1989, pp. 207-208) drew on their case expertise to question statistical measures of GNP per capita that suggested that China grew at a much faster rate than India before 1980. In turn, their alternative measure of GNP per capita refuted the idea that China’s better social performance was rooted in higher growth. As this example suggests, qualitative researchers are especially capable of assessing the meaning of indicators across diverse contexts. To cite another example, it is probably no accident that a qualitative researcher uncovered early programs of social provision in the United States that call into question the extent to which this country really was a social-spending laggard (Skocpol, 1992).

Finally, scholars who pursue qualitative research are less likely to find their work subject to “data-induced measurement error,” in which inaccurate, partial, or misleading secondary sources lead to the poor coding of variables (Bowman, Lehoucq, & Mahoney, 2005). The reason once again is the case-oriented nature of qualitative research: By learning a great deal about each of their cases, qualitative researchers tend to avoid simple coding errors that may be common in some large-N, statistical databases.

**Populations and Case Selection**

Almost all theories in the social sciences are designed to apply to some cases and not others; social science theories are bound by “scope conditions” that put limits on their generality. Specifying the scope of a theory accordingly must be viewed as a major part of theory development itself. In recent years, qualitative methodologists have begun to develop a set of best practices for researchers to follow when formulating this part of their theories (Collier & Mahoney, 1996; Mahoney & Goertz, 2004, 2006; Ragin, 2000).

These methodological writings in many ways were developed in response to criticisms pertaining to case selection in small-N research. It is therefore useful to briefly consider the critiques before turning to the positive contribution of qualitative methodology. Achen and Snidal (1989), Geddes (2003), and
King, Keohane, and Verba (1994) all suggest that selection bias in small-\(N\) studies produces serious error that can be avoided in good large-\(N\), quantitative research. The heart of the criticism is that qualitative studies “select on the dependent variable” and thereby fail to study samples with the full range of variation on this variable. Indeed, qualitative studies sometimes select only cases with the same value on the dependent variable, frequently an extreme value (e.g., exceptionally high growth rates), and thus from the perspective of mainstream statistical analysis cannot identify the causal factors that differentiate cases with this value from those without it. In specifically the field of comparative politics, Geddes has been most forceful in arguing that selection on the dependent variable is a major problem in well-known works on economic growth (Johnson, 1987), social revolution (Skocpol, 1979), and inflation (Hirschman, 1973).

One response of qualitative researchers to these concerns is to emphasize how within-case analysis, a technique of causal assessment discussed below, is not subject to the same kinds of selection issues that arise in large-\(N\) regression studies (see Collier, Mahoney, & Seawright, 2004). Likewise, qualitative methodologists may stress how the absence of variation on the dependent variable is not a methodological fallacy for studies that analyze necessary causes (Braumoeller & Goertz, 2000; Dion, 1998). These responses suggest that many concerns about selection bias in qualitative research are a fundamental misapplication of ideas from regression analysis to qualitative research.

However, another response, more relevant in the present context, involves the ways in which qualitative researchers constitute populations and establish the scope of their theories. Most basically, qualitative analysts often believe that samples drawn from large populations are heterogeneous vis-à-vis the theories that interest them. Qualitative researchers therefore choose to locate smaller populations of cases that exhibit sufficient similarity to be meaningfully compared to one other. In particular, they attempt to ensure that their populations of cases meet the assumptions of causal and conceptual homogeneity that are required for most modes of causal inference in the social sciences (Goertz & Mahoney, 2005; Ragin, 2000).

The absence of homogeneity (i.e., heterogeneity) produces instability in causal inference, and it can be generated in various ways. It may be that, in cases outside an initial context, unknown or unanalyzable variables are present that produce unstable causal inferences. Or the problem could be instability in the measurement of key variables across diverse contexts. Indeed, the very meaning of the units being analyzed (e.g., states) may change across temporal and spatial contexts, thereby generating heterogeneity and unstable estimates of causal effects.
The methodological challenge of establishing an appropriate scope therefore involves identifying a theory and population of cases that avoids omitted variable bias, incorrectly specified relationships among included independent variables, and unstable measurement across units and variables (Goertz & Mahoney, 2005). Although the case-intensive knowledge of qualitative research does not fully solve these problems, it does seem to help. For example, as already suggested, case-intensive knowledge helps analysts achieve measurement stability across units and thus enables them to make better decisions about where to draw the line when constituting homogeneous populations. Likewise, by virtue of developing contextualized knowledge about each of their cases, qualitative researchers are less likely to exclude key variables or misspecify the interrelations among included variables. On this point, it is worth noting that complex causal theories—such as path-dependent arguments (Pierson, 2004), two-level theories (Goertz & Mahoney, 2005), and theories of multiple causation (Ragin, 1987)—are relatively common in qualitative studies and relatively rare in quantitative studies (see also Coppedge, 1999; Hall, 2003). This is true because as qualitative researchers learn more about their cases, they come to increasingly appreciate the complexity of causal relationships in those cases. For instance, in the absence of her deep historical knowledge of France, Russia, and China, one can scarcely imagine Skocpol (1979) being able to hypothesize about the complex ways in which international pressure, dominant-class political leverage, and agrarian backwardness combine to produce state breakdown. Analogously, Collier and Collier (1991) surely would not have been able to hypothesize about the path-dependent linkages between labor incorporation and party-system regimes in Latin America without knowing a great deal about the modern histories of the Latin American countries.

The upshot of these considerations is that qualitative researchers often have good methodological reasons for treating a small number of cases as the full population about which they seek to generalize. If extending beyond this population risks introducing causal heterogeneity, they may be better off with the small sample. Moreover, qualitative researchers often are in a strong position to make suggestions about the kind of modifications that would be necessary to extend their causal theories to additional cases. For example, Skocpol (1979) suggests that extending her theory of social revolution beyond France, Russia, and China would require introducing the complexities of colonial legacies and taking into consideration the effects of economic dependence. However, she maintains that her structural perspective—which emphasizes state organization, international
pressures, and community solidarity networks among lower-class actors—would still be a highly fruitful approach for analysis.

Theory Testing

Theories in qualitative research may be both developed from and tested with a relatively small number of cases. From the perspective of quantitative methodology, this practice seems to violate a basic norm of scientific inference: The same cases cannot be used to both develop and test theory. Although one can legitimately question this norm if the goal of scientific research is to explain specific outcomes in particular cases (see Mahoney & Goertz, 2006), it is not necessary to do so here. For qualitative researchers need not use the same observations to develop and test their theories (e.g., Rueschemeyer, 2003). Instead, they can break cases apart into multiple observations, some of which are used to develop theory, others of which are used for theory testing. Rather than dwell on problems of data contamination, therefore, it is more productive to take a hard look at the actual methods of theory testing that are used in qualitative analysis.

Within-Case Analysis

Within-case analysis is a mode of causal inference in which researchers test hypotheses in light of multiple features of their cases. The technique has long been discussed in the field of qualitative methodology (e.g., Barton & Lazarsfeld, 1969; Campbell, 1975; George & McKeown, 1985), but in recent years there has been an intensification of effort to codify its procedures. Within-case analysis is featured in Collier et al.’s (2004, pp. 12, 252-258) discussion of a “causal-process observation,” defined as “an insight or piece of data that provides information about context or mechanism” (p. 252). Likewise, George and Bennett’s (2005, chap. 10) new discussion of process-tracing develops considerably the tools used “to identify the intervening causal process—the causal chain and causal mechanism—between an independent variable (or variables) and the outcome of the dependent variable” (p. 206). Although different authors use different labels to designate within-case analysis, all describe techniques in which hypotheses are evaluated by elucidating intervening processes and other observable implications of arguments.

To illustrate within-case analysis, an analogy between qualitative researchers and criminal detectives is instructive. Like a detective solving a crime,
qualitative comparative researchers use detailed fact collection and knowledge of general causal principles to explain outcomes (see Goldstone, 1997; McKeown, 1999). Not all pieces of evidence count equally. Some forms of evidence are like “smoking guns” that strongly suggest a theory is correct; other kinds of evidence are “air-tight alibis” that strongly suggest a theory is not correct. For qualitative researchers, a theory is often only one key observation away from being falsified. And yet, qualitative researchers often have certain kinds of evidence that suggest that the likelihood of theory falsification ever occurring is small.

In comparative politics, scholars use within-case analysis to locate the intervening mechanisms linking a hypothesized explanatory variable to an outcome. The identification of such mechanisms helps researchers avoid mistaking a spurious correlation for a causal relationship, even when only one or a small number of cases are analyzed. When clear mechanisms linking a presumed explanatory variable and outcome variable are identified, one’s confidence that the relationship really is causal is increased; if such mechanisms cannot be identified, one’s confidence about causality is challenged.

For example, Luebbert (1991, pp. 308-309) uses process tracing to eliminate the “Moore-Gerschenkron thesis,” which holds that fascist regimes result from the presence of labor-repressive landed elites who are able to draw substantial lower-class rural support for fascism. Although there is a matching between the presence or absence of a repressive landed elite and the presence or absence of fascism in the European cases, Luebbert suggests that the mechanisms through which this specific factor supposedly produces fascism are not supported by the historical record. Crucially, lower-class 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 (Luebbert, 1991, pp. 308-309). In short, despite the matching of macro variables, Luebbert rejects the Moore-Gerschenkron hypothesis because it is not validated by process tracing.

Qualitative researchers may use within-case analysis in ways other than the analysis of intervening processes. For example, a given hypothesis might suggest several auxiliary hypotheses that can be tested in conjunction with the main hypothesis of interest. Likewise, a given hypothesis might suggest other specific evidence that should be present if the hypothesis is correct. Such evidence may lend decisive support to a theory. An illustration is found in King et al.’s (1994, pp. 11-12) discussion of the meteorite collision theory of the extinction of dinosaurs. One observable implication
of this theory concerns the presence of iridium in a particular layer of the earth’s crust. The hypothesized iridium is not an intervening mechanism linking the meteorite collision to the dinosaur extinction. Rather, it is a separate piece of data that gives us confidence that, in fact, a meteorite collision did occur when it was hypothesized to occur. It is one of several observable implications of the theory (another would be the presence of a huge meteorite crater) that can be used to test it. These kinds of observable implications in effect are predictions about other outcomes that should be present; they are not intervening processes but deduced “facts” whose existence make a given hypothesis much more plausible. Even when small in number, they can greatly increase our conviction that a particular theory is valid.

An example from comparative politics is Marx’s (1998) study of racial orders in the United States, South Africa, and Brazil. Marx seeks to explain why the United States and South Africa experienced highly repressive racial domination, whereas Brazil was marked by more significant racial toleration. His cross-case argument emphasizes the role of intra-White divisions. Where Whites were divided, as in the United States and South Africa following the Civil War and Boer War, nationalist loyalty and White unity were forged through the construction of systems of racial domination that systematically excluded Blacks. Where no major intra-White cleavage developed, as in Brazil, Whites did not have to achieve unity through exclusion, and thus a much higher degree of racial harmony could develop.

Marx supports this argument using within-case evidence. Some of the evidence involves examining intervening mechanisms and thus follows the process-tracing approach discussed above. Yet additional evidence takes the form of facts that confirm implicit and explicit predictions about other things that should be true if this argument is valid. For instance, Marx suggests that if intra-White conflict really is decisive, efforts to enhance Black status should produce increased White conflict along the North-South fault line in the United States and between the British and Afrikaners in South Africa. By contrast, progressive racial reforms should not generate similar intra-White divisions in Brazil. His historical narrative then backs up this proposition. Likewise, if intra-White divisions indeed are the key, then Marx argues that we should see evidence that more progressive White factions view political stability as more important than racial equality. Again, he finds support for this argument. Overall, he suggests that it is highly unlikely that the auxiliary facts supportive of his argument are simply accidental; rather, he contends that they are symptoms of a valid main thesis.
Cross-Case Analysis

In qualitative research, hypotheses are also often evaluated through cross-case comparisons. Early discussions of cross-case techniques usually focused on Mill’s methods of agreement and difference (e.g., Skocpol & Somers, 1980) and Przeworski and Teune’s (1970) most similar and most different research designs. Over the years, many comparative studies have described their methodology as conforming to these techniques (e.g., Charrad, 2001; Collier & Collier, 1991; Orloff, 1993; Skocpol, 1979).

In more recent periods, however, three trends have pushed the discussion of cross-case methods in new directions. First, the focus on within-case analysis has served to deemphasize the importance of cross-case techniques of comparison. Most importantly, Collier et al.’s (2004) discussion in *Rethinking Social Inquiry* stresses within-case analysis—not cross-case analysis—as the distinctive source of leverage for causal inference in qualitative research. Second, the rediscovery of Mill’s method of concomitant variation (DeFelice, 1986; Mahoney, 1999) and its application to comparative studies has made it clear that some qualitative researchers pursue a kind of cross-case ordinal comparison that could be seen as intuitive regression. Studies that explore how small-\(N\) comparisons link to large-\(N\) comparisons also assume that qualitative researchers pursue intuitive regression (e.g., Coppedge, 1999; Lieberman, 2005). Third, techniques such as the methods of agreement and difference have been greatly extended and even superseded by the methodology of necessary and sufficient conditions, Boolean algebra, and fuzzy-set logic (e.g., Goertz & Starr, 2003; Ragin, 1987, 2000). These latter cross-case techniques are quite different from the procedures used in statistical analysis, and I discuss their prospects for shaping work in comparative politics here.

Qualitative methodologists have recast Mill’s methods of agreement and difference as tools for eliminating potential necessary and sufficient causes (see Dion, 1998; George & Bennett, 2005; Mahoney, 1999). 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. These methods are deterministic in that a single deviation from a hypothesized pattern of necessary or sufficient causation is enough to eliminate a given explanatory factor. Because of this deterministic character, which is quite controversial and unacceptable to some scholars, the methods provide a basis for systematically eliminating rival hypotheses when only a small number of cases are selected.
Critics of cross-case methods in qualitative research often show too little awareness of the methodological issues raised by necessary and sufficient causation (e.g., Goldthorpe, 1997; Lieberson, 1991). Moreover, they do not generally acknowledge that most limitations associated with Millian methods cannot be extended to Boolean algebra as codified in Ragin’s (1987) qualitative comparative analysis (QCA). Unlike the Millian methods, which focus on individual causes, QCA allows the analyst to treat several different combinations of variables as the causes of an outcome (Ragin, 1987). In particular, this methodology provides a logical basis for identifying combinations of causal factors that are sufficient for the occurrence of an outcome. Because several different combinations of factors may each be causally sufficient, the method further allows for multiple paths to the same outcome (which is sometimes called equifinality or multiple causation).

To best illustrate these methods, a simple example is needed, and Hicks, Misra, and Nah Ng’s (1995) analysis of the origins of a consolidated welfare state provides a good one. These authors focus on 15 developed countries, 8 of which consolidated welfare states by 1920. To explain welfare state consolidation, five dichotomous independent variables are examined: liberal government (LIB), Catholic government (CATH), patriarchal state (PAT), unitary democracy (UNI), and working-class mobilization (WORK). Through Boolean analysis, the authors discover that there were “three routes to the early consolidation of the welfare state”—(a) a Bismarckian route, (b) a liberal-labor route, and (c) a Catholic paternalistic route (p. 344). In particular, there were three combinations of variables sufficient for early welfare consolidation (capitalized variables are present, lowercase variables are absent): (a) cath*PAT*UNI*WORK, (b) LIB*cath*UNI*WORK, and (c) lib*CATH*PAT* UNI*WORK.

This basic approach to causation is often implicitly used in qualitative studies, even though results are not usually formally codified with Boolean algebra. The goal of such research is to identify necessary causes and combinations of factors that are sufficient for outcomes. In the Hicks et al. (1995) study, for example, unitary democracy and working-class mobilization are necessary causes that can join with different factors to produce a combination sufficient for a consolidated welfare state. With these results, it is difficult to generalize about the “average” effects of variables. For instance, consider the effect of Catholic governments on welfare state development. In the Bismarckian and lib-lab routes, these governments must be absent for welfare state consolidation. But in the Catholic paternalistic route, these governments must be present. One cannot therefore say whether Catholic governments help or hinder welfare states without saying something about the context (i.e.,
the other variable values) with which they appear. The average effect of these governments is not meaningful or interesting.

More recently, Ragin (2000) has introduced fuzzy sets as a means of continuously coding variables according to the degree to which they correspond to qualitative categories of interest. Fuzzy-set measurement is highly appropriate for the analysis of necessary and sufficient causation, including under probabilistic assumptions where different degrees of necessary or sufficient causation are considered. To employ the technique, the analyst must first measure all variables as fuzzy sets by coding them from 0 to 1. Causation is then assessed by comparing the values of cases on potential causal variables (including combinations of variables) with their values on outcome variables. With a necessary cause, fuzzy-membership scores on the outcome will be less than or equal to fuzzy-membership scores on the cause. By contrast, with a sufficient cause, fuzzy-membership scores on the cause will be less than or equal to fuzzy-membership scores on the outcome. Probabilistic benchmarks and significance tests can easily be used with fuzzy-set analysis, and a free software package that performs the mathematical operations is available (Ragin & Drass, 1999).

A growing number of studies in comparative politics have used Boolean and fuzzy-set techniques for testing hypotheses about necessary and sufficient causes. Examples of QCA works from major social science journals include Amenta’s (1996) study of New Deal social spending, Berg-Schlosser and DeMeur’s (1994) analysis of democracy in interwar Europe, Huber, Ragin, and Stephens’s (1993) work on the welfare state, Griffin, Botsko, Wahl, and Isaac’s (1991) study of trade union growth and decline, Mahoney’s (2003) work on long-run development in Latin America, and Wickham-Crowley’s (1991) study of guerrillas and revolutions. These works all identify complex patterns of causation, and it is debated how and if statistical methods could be used to analyze such patterns (see Bennett & Elman, 2006a; Braumoeller, 2003; Clark, Gilligan, & Golder, 2006).

Published QCA studies that analyze substantive issues also suggest a challenge facing scholars who use these techniques: balancing the technical apparatus of QCA with case-oriented research. In an article-length format, in particular, it is difficult to formally employ the technical side of QCA without sacrificing at least some of the case-based orientation of QCA. By contrast, book-length studies (e.g., Foran, 2005; Hicks, 1999; Wickham-Crowley, 1992) may more naturally combine both sides of the QCA methodology. The same could be said of within-case analysis: it may more easily be pursued in a book-length format where one has the space to thoroughly trace processes and discuss why particular observations lend critical support for or against a theory.
Conclusion

The future prospects of research that develops or uses qualitative methodology in comparative politics appear to be quite good. Qualitative methodologists and researchers have opportunities to publish their work with leading presses and in many top journals. Book publications long have tended to rely heavily on qualitative analysis, in part because the scholarly community often expects books to gather and analyze new data, which usually necessitates qualitative research. Although the editors at some top presses may currently be seeking to publish more work that employs formal models and/or statistical methods, the market for publishing excellent qualitative research on comparative politics in book form remains strong. The one cautionary note here concerns case study monographs that describe the contemporary politics of a particular country. Although these books continue to contribute much of the basic information we have about politics in specific countries, they may be increasingly hard to publish at leading presses unless they show a strong engagement with broader theory and an exceptionally high level of methodological rigor.

In terms of journals, recent changes at the American Political Science Review, the advent of Perspectives on Politics, and the continued commitment to diversity at World Politics appear to provide opportunities for qualitative researchers and methodologists in comparative politics to place their work in some of the discipline’s most prestigious outlets. Among more specialized journals, Comparative Politics publishes a high percentage of small-N, qualitative studies. Comparative Politics Studies, which is often now considered the flagship journal of the subfield, publishes research using a diverse portfolio of methodologies, including many qualitative articles. The revival of Studies in Comparative International Development as a leading journal in the field has been accompanied by the publication of a large number of excellent articles using and developing qualitative methods. When taken together, these publication opportunities suggest that qualitative research will continue to play a leading role in the field of comparative politics for the foreseeable future.

Notes

1. Skocpol (1979, p. 17) cites Wendell Phillips in making this declaration.
2. King, Keohane, and Verba’s description of how this theory was actually tested is not correct (see Waldner, in press), but the methodological point is still useful.
3. However, the American Journal of Political Science and the Journal of Politics—which are often regarded as prestigious disciplinary journals—publish virtually no qualitative articles in the field of comparative politics.
References


140 Comparative Political Studies


James Mahoney is an associate professor of political science and sociology at Northwestern University. His current research concerns colonialism and development in the non-European world. He is president of the Qualitative Methods Section of the American Political Science Association.