

Tools for Qualitative Research

Gerardo L. Munck

The late 1960s to mid-1970s was a major period of innovative writing on qualitative methodology and small-N research. Following an abatement of discussion, scholars again began to actively debate these aspects of methodology in the 1990s.¹ This new work has focused on a diverse set of issues, including case selec-

I would like to acknowledge the excellent and careful feedback I received from David Collier, Diana Kapiszewski, Sally Roever, and Jason Seawright, who generously commented on this article more than once. I am also grateful for the useful comments offered by Robert Adcock, Chad Atkinson, Ruth Berins Collier, Andrew Gould, Gary King, Alexander Kozhemiakin, James Kuklinski, James Mahoney, Sebastian Mazzuca, Richard Snyder, Jaroslav Tir, and Jay Verkuilen. Any errors that remain, of course, are my responsibility.

¹Some key works from the 1970s include: Smelser (1973; 1976; and see also 1968), Przeworski and Teune (1970), Sartori (1970), Lijphart (1971; 1975), and Eckstein (1975). Obviously, publication on comparative methodology did not cease during the late 1970s and the 1980s. See, for example, Skocpol and Somers (1980), Skocpol (1984), Sartori (1984), and Tilly (1984). This period, nonetheless, saw nothing similar to the current explosion of publications. Some of the most significant contributions to this methodological revival include: Ragin (1987; 1994; 2000), Ragin and Becker (1992), Sartori (1991), Geddes (1991), Collier and Mahon (1993), Collier and Mahoney (1996), Collier and Levitsky (1997), King, Keohane, and Verba (1994), Janoski and Hicks (1994),

tion, conceptual stretching, process tracing, the role of historical narratives in causal inference, and multiple conjunctural causation. Indeed, this new literature has addressed most issues that affect the conduct of research.²

While the contributions of a wide range of scholars are undeniable, it is equally true that the publication of one single book—Gary King, Robert O. Keohane, and Sidney Verba's *Designing Social Inquiry: Scientific Inference in Qualitative Research* (hereafter *DSI*)—has been a landmark event with an enormous impact on qualitative methods and research. *DSI*'s central message is that qualitative and quantitative research share a common logic of inference. Therefore, methodological lessons derived from one tradition can be applied fruitfully to the challenges faced by researchers in the other tradition. Unfortunately, *DSI* largely confines itself to applying tools of quantitative research to the problems of qualitative research, and undervalues the methodological insights and procedures that qualitative researchers bring to the table.

In fact, qualitative analysts have their own well-developed tools for addressing many tasks discussed by *DSI*. These tools certainly do not solve all of the problems faced by researchers, any more than quantitative tools do. Yet these qualitative tools deserve a central place within the standard repertoire of methodological practices. To balance the discussion, this chapter therefore considers some of the tools that qualitative researchers use in their efforts to produce valid social scientific inference. I consider specifically tools that qualitative researchers employ in five distinct steps in the research process.

The discussion below first shows how qualitative researchers seek to define the universe of cases to which their theories are deemed to apply, using contextually grounded analysis, typologies, and process tracing. Second, concerning case selection, I explore how qualitative researchers address the "many variables, small-N" problem. Qualitative analysts are often cautious about seeking to enhance inferential leverage by increasing the number of observations, recognizing that this practice may lead to problems of conceptual stretching and of causal heterogeneity. I discuss the approach of within-case analysis, and I stress that even though standard discussions of selection bias are clearly applicable to qualitative research, "no-variance" designs in qualitative research make an important contribution under some circumstances. I also show that qualitative researchers have long been concerned with the analytic leverage produced by different types of intentional case selection.

Tetlock and Belkin (1996), McDonald (1996), Mjoset, Engelstad, Brochmann, Kalleberg, and Leira (1997), Van Evera (1997), Bates, Greif, Levi, Rosenthal, and Weingast (1998), Peters (1998), J. S. Valenzuela (1998), Mahoney (1999; 2000a), Collier and Adcock (1999), Goldthorpe (2001), Abbott (2001), Mahoney and Rueschemeyer (2003), and George and Bennett (forthcoming).

²For an early effort at synthesis of this growing body of literature, see Collier (1993). See also Ragin, Berg-Schlosser, and de Meur (1996).

Third, regarding measurement and data collection, I discuss how qualitative researchers' concern with measurement validity may lead them to employ system-specific indicators and/or contextualized comparisons. I also explore the role of qualitative field research techniques such as in-depth interviews and participant observation. Fourth, I discuss qualitative procedures for causal assessment, with an emphasis on techniques for causal inference based on causal models other than the linear, additive model underlying most forms of regression analysis. I also consider the tools qualitative researchers use to distinguish systematic causal effects from causal effects produced by factors outside of the central hypothesis of concern, and I suggest why these tools are valuable.

In the fifth section, I go beyond *DSI*'s view of methodology as a set of tools primarily intended for addressing research questions that have already been formulated, and I consider the ongoing interaction among theory, hypotheses, and a given data set. Hypothesis testing is best seen as an iterative process that interacts with the development of theory, rather than as a process in which theory is more nearly treated as static. Table 7.1 provides an overview of research tools relevant to these several steps in the research process.³

Qualitative Methods: A Survey of Tools

Defining the Universe of Cases: Context, Typologies, and Process Tracing

A fundamental task in any research project is defining the universe of cases.⁴ Ideally, there is a close interaction between the investigator's understanding of this universe and choices about the theory that guides the study, the specific hypotheses to be investigated, the approach to measurement that is adopted, and the selection of cases for analysis. As investigators establish the fit between their hypotheses/models and the universe of cases, a standard concern is that, across the set of cases, the criteria of causal homogeneity⁵ and conditional independence should be met. Qualitative researchers have various tools for addressing these two issues.

To evaluate the assumption of causal homogeneity, in relation to a given set of cases and a particular explanatory model, qualitative researchers may turn this assumption into an initial hypothesis to be investigated in the course of research

⁴Many of these tools are, of course, not unique to qualitative investigation. The Point, rather, is that they are carefully and explicitly discussed in standard works on qualitative methodology.

⁵"Universe of cases" is a standard term in methodology; however, at certain points in the discussion below, it appears more natural to refer to this as the "domain of cases." This is sometimes called unit homogeneity.

Table 7.1. Tools for Qualitative Research*

Research Step	Task	Tool
DEFINING UNIVERSE OF CASES	Establish Causal Homogeneity	<i>Knowledge of context</i> Helps in assessing homogeneity of causal processes.
		<i>Ragin's QCA and critical juncture/path dependency frameworks.</i> Qualitative Comparative Analysis and these other frameworks point to additional variables that explain and potentially overcome causal heterogeneity.
		<i>Within-case analysis.</i> Evaluates causal processes within cases.
CASE SELECTION	Establish Conditional Independence	<i>Scope restrictions.</i> Specify appropriate domains of comparison.
		<i>Typologies.</i> Serve to identify multiple domains of causal homogeneity.
		<i>Within-case analysis, process tracing.</i> Help identify reciprocal causation. These tools, especially when focused on a sequence of causal steps, serve to test for reciprocal causation as part of the theory.
CASE SELECTION	Add Observations without Overextending the Analysis	<i>Reconceptualization.</i> Addresses conceptual stretching through mutual fine-tuning of concepts and case selection.
		<i>Addressing causal homogeneity and conditional independence.</i> Help in dealing with problems of overextension.
		<i>No-variance designs.</i> Facilitate close examination of causal mechanisms and yield descriptive insight into novel political phenomena.
CASE SELECTION	Select Cases Nonrandomly	<i>Matching cases on independent variables.</i> Serves the same purpose as statistical control.
		<i>Selecting sharply contrasting cases.</i> May permit stronger tests of hypotheses through focus on diverse contexts. High variability specifically on rival explanations may yield more leverage in test of theory.
Research Step	Task	Tool
MEASUREMENT AND DATA COLLECTION	Increase Measurement Validity	<i>System-specific indicators.</i> Use of distinct indicators in different settings.
		<i>Contextualized comparison.</i> Achieves analytic equivalence across contexts by focusing on phenomena that, in concrete terms, appear distinct.
		<i>In-depth interviews, participant observation, qualitative content analysis.</i> Yield data of greater depth compared with quantitative data sets.
CAUSAL ASSESSMENT	Collect Data	<i>Crucial experiments, crucial case studies.</i> Focus on cases that provide strong tests of a deterministic hypothesis.
		<i>Testing deterministic hypotheses against probabilistic alternatives.</i> Serves to bridge these alternative causal models.
		<i>Boolean algebra.</i> Evaluates deterministic causes.
CAUSAL ASSESSMENT	Assess Deterministic Causation	<i>Critical juncture and path dependence frameworks.</i> Offer a systematized approach to assessing historical causation.
		<i>Within-case control.</i> Serves to isolate analytically relevant components of phenomena and provides a substitute for statistical control, based on within-case analysis and process tracing.
ITERATED REFINEMENT OF HYPOTHESES AND THEORY	Assess Historical Causation	<i>Hypothesis testing and refinement of concepts.</i> Reframe and sharpen the analysis throughout the research cycle.
		<i>Separate Systematic vs. Random Components</i>
ITERATED REFINEMENT OF HYPOTHESES AND THEORY	Inductive Learning from Data	<i>Case studies.</i> Different types of case studies—heuristic, hypothesis-generating, disciplined-configurative, and deviant case studies—as well as no-variance designs, serve to generate new explanations.
		<i>Identify New or Alternative Explanatory Factors</i>

*Many of these tools are, of course, not unique to qualitative investigation. The point, rather, is that they are carefully and explicitly discussed in standard works on qualitative methodology

(Ragin, Berg-Schlosser and de Meur 1996: 752-53; see also Ragin 125-28 this volume). Although qualitative analysts have many procedures for assessing causal homogeneity, three deserve special attention here. First, researchers often use close knowledge of the cultural, historical, and political context to evaluate whether the causal processes identified in the hypothesis have the same form and significance across the various cases. Within the comparative-historical research community, this process corresponds to the effort to find the boundaries of causal arguments that is a central concern of what Skocpol and Somers (1980: 178—81) call the "contrast of contexts" approach to historical comparison.

Second, qualitative researchers may seek to *achieve* causal homogeneity by considering the various factors that could produce heterogeneity and conceptualizing them as additional variables to be included in the analysis. If, in the course of the analysis, these variables prove unimportant, they are discarded; otherwise they ultimately form part of the substantive explanation produced by the study. This process is perhaps most widely known in the formalized, Boolean-algebraic version created by Ragin (1987), which he calls Qualitative Comparative Analysis (QCA). However, qualitative researchers commonly apply informal versions of the same approach. For example, analyses that employ the frameworks of critical junctures (Collier and Collier 1991: chap. 1) or path dependency (Pierson 2000) follow this technique. These approaches typically identify variables that place countries (or other cases) on different paths or trajectories of change. Such trajectories often involve causal processes that work themselves out in contrasting ways within different groups of cases. The critical juncture can thus be understood as an event that explains subsequent causal heterogeneity. In this specific sense, the causal heterogeneity is explained and thereby effectively overcome.

Third, qualitative researchers assess causal homogeneity by applying different forms of within-case analysis. They examine detailed evidence about the causal process that produced the outcome of concern. For example, if the focus is on institutional decision making, qualitative researchers may analyze records of the conversations and thought processes involved in that decision making, using what Alexander George and Timothy McKeown (1985: 34—41) describe as process tracing. More generally, analysts search for evidence about the causal mechanisms that would give plausibility to the hypotheses they are testing. If this evidence suggests that a similar mechanism produced or prevented the outcome in each case, this constitutes evidence for causal homogeneity.

These procedures help scholars make carefully calibrated statements about the appropriate universe of cases, involving "scope restrictions" (Walker and Cohen 1985) that delimit the domain to which the argument applies. For example, Theda Skocpol (1979: 40-42, 287-92; 1994: 4-7) argues that it would be a mistake to apply her original theory of revolution directly to twentieth-century revolutions. This is because a central feature of the cases she studied, the presence of agrarian-

bureaucratic monarchies that had not experienced colonial domination, is simply not present in most twentieth-century revolutions. Although recognition that theories are bounded in this manner is also found in quantitative research, qualitative researchers have generally been more sensitive to this issue.

An alternative approach to assessing causal homogeneity is to identify multiple domains, within each of which the analyst finds causal homogeneity and between which there is causal heterogeneity. Researchers routinely present such findings in the form of *typologies*. This use of typologies merits particular emphasis here, given that *DSI* dismisses them as a research tool of limited value (48). Yet, as George and McKeown (1985: 28-29, 45) argue, typologies can play a valuable role in defining the universe of cases that can productively be compared (see also Stinchcombe 1968: 43⁷; Ragin 1987: 20, 149).

For instance, establishing typologies of political regimes has been very useful in helping scholars delimit domains of cases. Perhaps the most influential set of typologies of regimes is that associated with Juan Linz (1964; 1975). Linz and others working within his general framework distinguish, for example, among democratic, authoritarian, totalitarian, post-totalitarian, military, one-party, and sultanistic regimes. This family of typologies has played a key role in helping analysts of regime change identify universes of cases within which causal processes are seen as working in similar ways. For example, Linz and Stepan (1996: 55-64) theorize that regime type, defined according to the categories noted above, affects the probability and nature of regime change. Transitions from a given type of regime may tend to have dynamics and explanations that are similar to one another, but different in comparison to transitions from other regime types. Geddes (1999) argues that the type of regime that existed prior to the transition—one-party, military, or personalistic/sultanistic—defines domains of cases within which the causal story of transition involves different independent variables. She thereby specifies domains of causal homogeneity. Thus, typologies can play a central role in developing statements about the scope of theories.⁶

Qualitative researchers also address the criterion of conditional independence, which includes the challenges of avoiding endogeneity (i.e., a situation in which the values of the explanatory variables are caused by the dependent variables) and of including all-important explanatory variables. Within-case analysis is again valuable here, in that it encourages researchers to identify and analyze the temporal sequence through which hypothesized explanatory variables affect outcomes.

⁶On efforts to ensure causal homogeneity, see also the discussion of "frames of comparison" and "contrast space" in Collier and Mahoney (1996: 66-69) and of positive and negative cases in Ragin (128-33 this volume). These various suggestions are still in need of refinement. Nonetheless, they are certainly worth pursuing, especially given Bartels's (74 this volume; see also 1996) argument that quantitative methodologists have still not dealt with this problem adequately, even though it may be possible to address causal heterogeneity with a complex regression model.

Within-case analysis privileges evidence about causal mechanisms, pushing researchers to ask whether change in the independent variables in fact preceded change in the dependent variable and, more significantly, by what process change in the independent variables produced the outcome. This process of studying sequences of change may also alert qualitative researchers to important missing variables, thereby addressing another aspect of the conditional independence assumption. A focus on sequences and changes over time is by no means unique to qualitative research; quantitative researchers obviously analyze time-series data. The point here is simply that qualitative researchers likewise have tools for this type of analysis.

Of course, in many studies endogeneity is impossible to avoid. In these situations, qualitative researchers may seek to focus explicitly on the reciprocal interactions among relevant variables and make inferences about the several causal links involved. This focus is found, for example, in studies that analyze "virtuous" or "vicious" cycles of political and economic events and of policy change,⁷ as well as in studies of the dynamic interaction among leaders or other political actors.⁸

Case Selection: Dilemmas of Increasing the Number of Observations

A recurring piece of advice regarding case selection is to increase inferential leverage by adding new observations beyond those previously studied. This procedure is recommended repeatedly by *DSI*,⁹ and it is extensively discussed in standard treatments of qualitative methodology (Lijphart 1971: 686; Smelser 1976: 198-202). *AST*'s advice that qualitative researchers increase the number of observations drawn from within the cases already being analyzed (24, 47, 120, 217-28) corresponds to a standard practice among qualitative researchers.¹⁰

However, three concerns must be raised about increasing the number of observations. First, it may be "neither feasible nor necessarily desirable" (Ragin, Berg-Schlosser, and de Meur 1996: 752), and in many ways this advice amounts to little more than saying that "qualitative researchers are inevitably handicapped" and that they should "not be 'small-N' researchers" after all (Brady 55 this volume; see also McKeown 145-46 this volume)."

⁷See, for example, Kahler (1985: 477-78); Doner (1992: 410); Kapstein (1992: 271); Pierson (1993: passim); and Costigliola (1995: 108-9).

⁸See, for example, Stepan (1978), Higley and Gunther (1992), or Linz and Stepan (1996: 87-115).

⁹*DSI* 52, 67, 99, 116-20, 178-79, 213-17, 228.

¹⁰Smelser (1973: 77-80; 1976: 217-18), Campbell (1975), George and McKeown (1985), Collier and Mahoney (1996: 70).

¹¹Like Lijphart (1971: 685), the authors of *DSI* operate with the assumption that we would always be better off using quantitative methods, and that small-N research and the

Second, if a qualitative researcher does choose to study more observations, *DSI*'s advice fails to recognize the problem of conceptual stretching that can arise when new cases are studied or when the use of within-case analysis brings about a shift in the unit of analysis (Ragin 125-28 this volume).¹² Conceptual stretching is the problem of taking concepts that validly apply to a given set of cases and extending them to a domain where they do not fit. While some might see this problem as an insurmountable obstacle that would simply make comparative analysis untenable, the pioneering work on conceptual stretching by Sartori (1970; 1984; 1991), recently reworked and refined by Collier and collaborators,¹³ has sought to spell out procedures to guide the reconceptualization that may be needed to avoid conceptual stretching. Thus, insights developed by qualitative methodologists go considerably further than *DSI* in offering practical suggestions for dealing with this fundamental methodological challenge.

Third, efforts to increase inferential leverage by adding new cases may raise problems of causal heterogeneity. As discussed above, qualitative researchers are often hesitant to assume that causal homogeneity holds across a given range of cases, and they devote considerable attention to testing for heterogeneity. Extending an analysis beyond the domain for which causal homogeneity has been established requires researchers to choose between: (a) simply assuming that causal homogeneity holds among the new cases; or, (b) intensively testing each new case for causal homogeneity and including only those cases that pass the test, a process that may demand resources that could be better devoted to intensive analysis of the original set of cases.

DSI (116, 126-32, 135) gives considerable attention to the problem of selection bias.¹⁴ The authors present the standard argument that selecting on the dependent variable can yield cases that are skewed to the high or low end of the distribution on that variable, with the likely consequence of biasing estimates of causal effects. Qualitative researchers are advised, as a first solution, to select their cases on the independent variable. This approach eliminates a significant source of selection bias, although *DSI* (129, 141, 147-49) emphasizes that in selecting on the independent variable, scholars should seek sufficient variation. Alternatively,

comparative method should only be used as a backup option, when quantitative methods cannot be used.

However, this problem generally does not emerge in within-case analysis that generates the causal-process observations discussed in detail below (chap. 11), as opposed to data-set observations.

"Collier and Mahon (1993) and, since the publication of *DSI*, Collier (1995b), and Collier and Levitsky (1997).

This issue has been among the most debated aspects of *DSI*. See Collier, Mahoney, and Seawright (chap. 6, this volume), the exchange between Rogowski (77-82 this volume) and King, Keohane, and Verba (188-91 this volume), and Dion (1998).

scholars can select on the dependent variable, but here again it is essential to ensure an appropriate range of variation.

Yet *DSPs* advice concerning selection bias rests on the premise that causal inference requires the analysis of covariation between independent and dependent variables, a premise that can often be problematic (Lieberson 1985: 90-91; Ragin 1994: 107, 145-48). Because qualitative work often assesses causal effects through an analysis of covariation," *DSPs* insistence that studies include variation on both the explanatory and the dependent variable is, of course, relevant to qualitative researchers. However, many qualitative researchers make causal inferences by focusing attention centrally on processes and decisions *within* cases. While such analysis is certainly framed by at least implicit comparison with other cases, it is a different research strategy from that of explicit and systematic comparison. If this close analysis of processes and decisions focuses only on cases where the overall outcome being explained (e.g., war or revolution) has occurred, then it may be called a *no-variance design*. Qualitative researchers see such studies as making a key contribution in the research process, helping to generate the kind of insights into causal mechanisms without which the analysis of covariation is incomplete. This kind of design can be valuable for gaining descriptive insight into a political phenomenon about which researchers have little prior knowledge.

A great deal of methodological attention has been paid to research designs in which the analyst intentionally selects cases that do not vary on the dependent variable. However, these research designs should be situated in relation to the broad range of intentional case selection strategies that qualitative researchers routinely employ. Cases matched on independent variables may be selected, for example, to control for the effects of these explanatory factors. Sharply contrasting cases may be selected to explore the hypothesis that a given cause produces an outcome across various domains. These designs correspond to the standard procedures for analyzing matching and contrasting cases discussed by J. S. Mill (1974b [1843]) and by Przeworski and Teune (1970: 32-39). Cases that exhibit substantial variability on important rival explanations may be selected to provide a difficult test for a theory (Eckstein 1975: 113-32). These three approaches to intentional case selection provide qualitative researchers with valuable leverage in testing their hypotheses.

¹⁵It bears emphasizing, as Collier and Mahoney (1996: 75-80) argue, that many studies that are seen to lack variance on the dependent variable actually do exhibit variance. Part of the reason for this misperception is the fact that analysts fail to see how the study of cases over time naturally introduces variance on the dependent variable. *DSI* (129) does not appear to appreciate the significance of the longitudinal dimension of much comparative research, as the discussion of Skocpol's work on revolution demonstrates.

Measurement and Data Collection

With regard to measurement, *DSPs* lack of attention to standard methodological texts on—and the established practices of—qualitative research is again apparent in its overly brief discussion of measurement validity. *DSI* (25, 153) is on solid ground in calling for qualitative researchers to maximize the validity of their measurements. However, the book does little to incorporate prior work by comparativists who have grappled with the problem of validity,¹⁶ or to acknowledge the difficulty of developing equivalent indicators across different cases. For qualitative researchers, a key aspect of the problem is, simply, that just as words can take on different meanings when used in different contexts, indicators can also measure different things in different contexts. To take a traditional example, while the magnitude of economic activity can be measured quite accurately in monetary terms in western societies, money is an incomplete indicator in less developed societies that are not fully monetized (Smelser 1973: 69). More recently, concerns with this indicator arise due to the magnitude of the extralegal or underground economy in many developed countries. Thus, a researcher cannot assume that the same indicator will be a valid measure of a concept across different cases and time periods.

Qualitative researchers, for the most part, have not been self-conscious about ensuring measurement validity. Nonetheless, as Collier (1998a: 5) suggests, the close familiarity that qualitative researchers tend to have with their cases has allowed them to implicitly follow the long-standing advice of Przeworski and Teune (1970: chap. 6) to construct "system-specific indicators" as opposed to "common indicators" (see also Verba 1971; Zelditch 1971). More recent recommendations for tackling this problem have been offered by Locke and Thelen (1995), who urge scholars to carry out "contextualized comparison."¹⁷ Thus, *DSPs* discussion can be criticized on two grounds. First, it ignores key earlier literature, merely making the general argument that researchers should ensure the validity of their measurements (25, 153) and draw upon their knowledge of context (43), but failing to focus on specific procedures for accomplishing this in comparative research. Second, *DSI* fails to note that the sensitivity to context that researchers bring to small-N studies gives them an alternative form of leverage in dealing with issues

¹⁶Early discussions of measurement equivalence that draw on both the quantitative and qualitative traditions include: Przeworski and Teune (1970), Zelditch (1971), and Warwick and Osherson (1973: 14-28). See also Smelser (1976: 174-93).

¹⁷Considering the study of labor politics and economic restructuring, Locke and Thelen (1995) argue that a researcher should not simply focus, for example, on disputes over wages. Instead, a researcher should search for those points where conflicts emerge, "which might vary from case to case. Thus, to ensure the equivalence of measurements one might have to focus on conflicts over wages in one case, over employment in another, and over working hours in yet another.

of validity, compared to large-N researchers. An important reason for choosing a small N is thus simply ignored.

With regard to data collection, qualitative researchers employ intensive methods that produce richer, more multifaceted information than is contained in most quantitative data sets. In-depth interviews provide qualitative researchers with a great deal of valuable evidence. In such interviews, informants not only answer the specific, prepared questions that the researcher poses, but often offer their own more nuanced responses and unprompted insights. For these reasons, such interviews do not constitute a single "data point" in any normal sense; rather, they are a complex array of data, different parts of which can be used to support or undermine a theory. Other common qualitative practices such as participant observation and content analysis produce data that has similar "depth."

Causal Assessment in Cross-Case and Within-Case Designs

Much of quantitative researchers' treatment of causal assessment is essentially based on a standard regression model. This model tends to assume, as a default position, that causal effects are uniform across cases and operate in a probabilistic fashion (Abbott 1988; Abbott 1992: 432-34). Qualitative researchers, by contrast, have frequently employed different models of causation, and they utilize a variety of tools appropriate to these models.

First, qualitative researchers sometimes use a deterministic, as opposed to a probabilistic, model of causation (Ragin and Zaret 1983: 743-44; Ragin 1987: 15-16, 39-40, 52; Ragin 135-38 this volume), and have designed procedures for assessing this model. A deterministic understanding of causation, which allows the analyst to reject a potential explanatory factor on the basis of a single deviation from an overall pattern (Dion 1998: 128), is implicit in arguments that even single case studies can be used to test theories. Well-known examples include Lijphart's (1971: 692) "crucial experiments" and Eckstein's (1975: 113-32) "crucial case studies" (see also Rogowski 77-82 this volume). More recent discussions have creatively focused on the problem of testing the hypothesis of deterministic causation against the alternative hypothesis of probabilistic causation (Dion 1998; Ragin 2000; Braumoeller and Goertz 2000; Seawright 2002a,b).

Second, additional tools employed by qualitative researchers for testing alternative models of causation include Ragin's (1987, 2000) Qualitative Comparative Analysis (see above), which is used to test multiple, conjunctural causes; the use of Mill's methods jointly with process tracing to test what Stinchcombe (1968: 101-29) designates as "historical" as opposed to "constant" causes; and the closely related analytic procedures offered by the growing literature on critical

junctures and path dependence.¹⁸ Once again, quantitative researchers likewise have procedures for assessing these specific models of causation;¹⁹ I would merely stress that qualitative researchers have a long history of working with such models.

Third, as Seawright and Mazzuca argue, through the procedure they call "within-case control," qualitative researchers have a distinctive means of addressing an aspect of descriptive inference that *DSI* (56-61) emphasizes strongly: distinguishing between outcomes that are systematic with respect to a given theory and outcomes that are random with respect to that theory, or that are better treated as the result of different processes.²⁰ The idea of separating the systematic component of a phenomenon from the random component, summarized in chapter 2 of the present volume, is one of the three basic components in *DSI*'s account of descriptive inference.

Though the reason for making this distinction may be unclear to some researchers, it is in fact valuable in qualitative analysis for two closely linked reasons. First, in qualitative research it is difficult to introduce control variables. Hence, disaggregating the dependent variable by removing variation that is caused by factors other than those central to the explanatory model is a way of meeting the "other things being equal" criterion necessary for causal inference, and thereby achieving within-case control. Second, some causal factors are genuinely outside of the researcher's explanatory framework, and removing variance that results from these factors permits better inference about the aspects of social phenomena that are of greatest theoretical interest. For example, it may be interesting for a social movements scholar to learn that the intensity of some urban riots in the United States during the summer of 1968 was increased by hot weather, but this

For valuable discussions of methodological issues that arise in developing critical juncture/path-dependent models, see Collier and Collier (1991: chap. 1), Jackson (1996: 722-26, 730-45), Pierson (2000), and Mahoney (2000b). For discussions of critical juncture models in research on party systems, regime change, and economic transformations, see Lipset and Rokkan (1967), Collier and Collier (1991), Stark (1992), and Ekiert (1996).

Beyond the distinctive issues raised by deterministic, multiple, conjunctural, and historical causes, a significant challenge concerns the assessment of models of asymmetrical (Lieberson 1985: chap. 4) and cumulative causation (Stinchcombe 1978: 61-70). See also Zuckerman (1997). I would stress that the need to assess this range of causal models is not a point that divides quantitative and qualitative researchers. Thus, it is noteworthy that quantitative methodologists have also sought to devise tools to assess necessary and sufficient causes (Braumoeller and Goertz 2000), models of multiple causal paths (Braumoeller 1999), and path dependent causes (Jackson 1996: 730-45), and, more generally, have sought to fashion quantitative methods more suited to historically oriented analysis (Griffin and van der Linden 1999).

Jason Seawright and Sebastian Mazzuca, personal communication.

scholar might well want to remove this aspect of the variance in the outcome, to permit a more direct test of social and political hypotheses.

Qualitative researchers can achieve within-case control by closely examining the causal process and separating out distinct components of the variance being explained. Within-case analysis helps researchers assess to what degree the mechanism hypothesized by a theory was present among all the cases under study. Researchers can thus make inferences not only about the extent to which the hypothesized cause was found across cases, but also about the extent to which that cause produced the outcome for each case. For deviant cases, that is, cases that do not follow the causal pattern predicted by the theory, within-case analysis gives qualitative researchers an opportunity to discover the processes that caused the case to diverge from the hypothesized outcome. These processes may involve variables quite unrelated to the main hypothesis, and therefore may be seen as random with respect to that hypothesis. However, in qualitative research the variance associated with these processes is not automatically separated out, as it is in regression analysis. Rather, the researcher must carefully consider evidence about the nature of each "random" process in order to eliminate from the dependent variable the variance associated with that process.

The value of separating the systematic and the random component through within-case control may be illustrated by an example. Thomas Ertman's (1997) analysis of early-modern state building hypothesizes that the interaction of (a) the type of local government during the first period of state-building, with (b) the timing of increases in geopolitical competition, strongly influences the kind of regime and state that emerge. He tests this hypothesis against the historical experience of Europe and finds that most countries fit his predictions. Denmark, however, is a major exception. In Denmark, sustained geopolitical competition began relatively late and local government at the beginning of the state-building period was generally participatory (305-6), which should have led the country to develop "patrimonial constitutionalism." But in fact, it developed "bureaucratic absolutism." Ertman carefully explores the process through which Denmark came to have a bureaucratic absolutist state and finds that Denmark had the early marks of a patrimonial constitutionalist state. However, the country was pushed off this developmental path by the influence of German knights, who entered Denmark and brought with them German institutions of local government (307). Ertman then traces the causal process through which these imported institutions pushed Denmark to develop bureaucratic absolutism (307-11), concluding that this development was caused by a factor well outside his explanatory framework. Ertman makes a parallel argument for Sweden (311-14), and summarizes his overall interpretation of these cases by stating that:

In both Sweden and Denmark, the two factors highlighted throughout this book also operated, broadly speaking, in the manner expected—Yet in both cases contingent historical circumstances intervened to shunt these states off the path

leading to noble dominance and patrimonial constitutionalism and onto rather different roads. (Ertman 1997: 316)

This conclusion could be misunderstood as an inappropriate attempt to discard information that runs counter to the main hypothesis. A better way of thinking about this, as we have emphasized, is to see it as analogous to the initiative in quantitative research of introducing a control variable. Adding a control variable in effect poses the question: other things being equal, does the main hypothesis in fact explain part of the outcome? Through within-case control, qualitative researchers have a means of addressing this question.

Beyond Strict Hypothesis Testing: Theory Generation, Reformulation, and the Iterated Assessment of Hypotheses

Quantitative methodologists often take a relatively strict view of hypothesis testing, issuing warnings against data mining and against testing a given hypothesis with the data used to generate it. Qualitative methodologists, on the other hand, point to opportunities for moving beyond strict hypothesis testing by engaging in the ongoing refinement of concepts, the iterated fine-tuning of hypotheses, and the use of specially targeted case studies that appear likely to suggest new hypotheses and theoretical ideas.

DSI undervalues the contribution to theory development and reformulation that is made by ongoing interaction with the data. *DSI*'s cautionary remarks about reformulating the theory after analyzing the data (21-22) and about data mining (174) are unduly restrictive. Theory reformulation that occurs after looking at the data is critical because it allows social scientists to learn from their research. Indeed, it would be an important constraint on the accumulation of knowledge if analysts did not routinely revise their explanations of a set of cases and then test the new explanation—if need be, with the same set of data. The concerns with contextual specificity discussed above may convince the researcher that moving beyond this initial set of cases is not analytically productive. Of course, careless revisions of theory should be considered suspect, yet it is vital to recognize the legitimacy of efforts to inductively reformulate theory by carefully incorporating insights drawn from research findings.

With regard to refining concepts, Ragin (125-28, 130-33 this volume) suggests that an ongoing process of concept formation should be intimately interconnected with the analysis of positive and negative cases that exemplify the variation of interest. This does not occur merely at the onset of a study, but is a process that continues throughout the study. More generally, scholars frequently refine their

variables, often through disaggregation, in order to more adequately capture the ideas involved in the hypotheses they are testing.²¹

Qualitative researchers routinely build on their in-depth knowledge of cases to gain further insights about causal processes (Collier 1999), which among other things can improve causal inference by suggesting important missing variables. To do this, qualitative researchers rely on a spectrum of case-oriented research designs, such as Lijphart's (1971: 691–93) "hypothesis-generating" case study, which corresponds to what Eckstein (1975) calls a "heuristic" case study; Eckstein's "disciplined-configurative" case study; and the "no-variance" small-N designs discussed above. Lijphart's "deviant" case-study design, like these other approaches, can play a central and creative role in suggesting further hypotheses.

The core point, as Ragin (125–28, 135–38 this volume) states, is that researchers should not treat tests of causal hypotheses as the endpoint of a study, but rather as an ongoing activity that should be closely intertwined with these other components of the research process.

Conclusion

This chapter argues that just as quantitative researchers can draw upon a relatively standardized set of methodological procedures, so qualitative scholars also have well-developed procedures—which in fact address every step in the research process. The problem is not that qualitative researchers lack tools to conduct their research, but rather that these tools have not been adequately systematized. The goal of this chapter has been to formulate them more systematically (see again table 7.1).

Although qualitative researchers can take considerable satisfaction in this set of tools, the contributions of qualitative methodology should not be overstated. As Bartels (74 this volume) suggests, part of the problem with *DSI* is that its authors "promise a good deal more than ... [they] could possibly deliver given the current state of political methodology" (see also Brady 55–56 this volume; Jackson 1996: 742–45). Correspondingly, even though *DSI* persistently undervalues the contributions of qualitative methodologists (McKeown 145–46 this volume), qualitative researchers should not try to correct this imbalance by overselling their own ap-

proach. Substantively oriented research will be advanced most effectively to the extent that a more meaningful dialogue between quantitative and qualitative researchers is established, and the strengths of alternative methods are brought to bear on interesting questions of political analysis.

²¹Skocpol's (1979) research on social revolution exemplifies this approach. She disaggregates her dependent variable into two parts—state breakdown and peasant uprising—a decision that allows her to build her argument around two distinct, though interrelated outcomes. This allows her to focus more clearly on the mechanisms that generate these distinct outcomes. In addition, she is able to avoid potential confusion by showing how certain variables (e.g., international pressures) are used to explain state breakdown and not (at least not directly) peasant uprising. Finally, this approach allows Skocpol (1994) to integrate her findings as well as those of other researchers in the context of a general framework.