The Qualitative Foundations of Political Science Methodology

George Thomas

David Collier and Henry Brady, eds., *Rethinking Social Inquiry: Diverse Tools, Shared Standards* (Lanham: Rowman and Littlefield, 2004).

Colin Elman and Miriam Fendius Elman, eds., *Bridges and Boundaries: Historians, Political Scientists, and the Study of International Relations* (Cambridge: MIT Press, 2001).

Barbara Geddes, *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics* (Ann Arbor: University of Michigan Press, 2003).

John Gerring, *Social Science Methodology: A Criterial Framework* (New York: Cambridge University Press, 2001).

Charles Ragin, *Fuzzy-Set Social Science* (Chicago: University of Chicago Press, 2000).

The last decade has seen a contentious dialogue between quantitative and qualitative scholars over the nature of political science methodology. Even so, there has often been a consensus that quantitative and qualitative research share a "unified logic of inference;" that the differences between these "traditions are only stylistic and are methodologically and substantively unimportant."[1] All of the books under review here share these convictions. Yet the most remarkable feature of these works taken as a whole —and the focus of this review essay—is the more capacious view of the scientific enterprise on display. John Gerring's *Social Science Methodology,* David Collier and Henry Brady's *Rethinking Social Inquiry,* and Charles Ragin's *Fuzzy-Set Social Science* all focus on aspects of the scientific process beyond the testing of hypotheses—science being "a systematic, rigorous, evidence-based, generalizing, nonsubjective, and cumulative" way of discovering the truth about the world (Gerring, p. xv). If science is the systematic gathering of knowledge, testing hypotheses—the central concern of statistical inference—is an important part of this. But it is only one part. Before we can turn to testing hypotheses, we must be clear about concepts, theories, and cases. And here both Barbara Geddes's *Paradigms and Sand Castles* and the Elmans' *Bridges and Boundaries* complement the other works by attending closely to these issues even when the larger goal remains the *testing* of theory.

Each of these five books argues that social science methodology (1) requires systematic and continuous concept formation and refinement, (2) employs empirical evidence not only to confirm but also to develop and explore theories, and (3) must come to terms with causal complexity. Taken together, they provide support for a provocative possibility: if social science has a unified logic, it is found in approaches traditionally associated with qualitative methods rather than statistical inference.

Getting at the Arguments

Whether scholars are engaged in quantitative or qualitative research, or following rational choice or modeling approaches, we should share standards of evaluation and a common language. But there are important disagreements over what these shared standards are, as well as important differences in language and terminology. Geddes, while writing for comparativists of a qualitative bent, adheres to the foundational logic of statistical inference, as do a few authors in *Bridges and Boundaries.*[2] Gerring, Collier and Brady, and Ragin reject the notion that this quantitative template can serve as the foundation for political science as a whole, as do most authors in *Bridges and Boundaries.* Naturally, this division leads to important differences over what constitutes valid research and how it should be carried out.

George Thomas is assistant professor of political science at Williams College (gthomas@williams.edu). *The author thanks the anonymous reviewers for* Perspectives on Politics, *Mark Blyth, Brian Taylor, and, especially, Craig Thomas for helpful comments.*

https://doi.org/10.1017/S1537592705050486 Published online by Cambridge University Press

Research design and the science of political science

Geddes's *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics* is a high-level introduction to research design that offers a very lucid and detailed explication of what might be dubbed the conventional wisdom (for example, King, Keohane, and Verba, 1994). That is, the underlying framework is drawn from statistical inference. Geddes's detailed and easy-to-read chapters focus on how the questions, cases, evidence, and approach one chooses in carrying out one's research affect the answers one gets. She suggests that big questions might be broken down into smaller questions that can more readily be tested. And testing theory against empirical evidence is primarily what Geddes has in mind when she speaks of theory building. In doing so, we should posit clear and falsifiable hypotheses deduced from our theory. We should then test these hypotheses against the universe of cases to which they apply. Most importantly, we must test our hypotheses against cases that were *not* part of the inductive process from which the argument was initially proposed. Ideally, this should include as many observations against each hypothesis as possible. In fact, Geddes argues that we unfortunately cling to the details of our knowledge and the cases we know at the expense of building theory. Compartmentalization based on substantive knowledge (and particularly based on geographic regions) has no theoretical justification. "A carefully constructed explanatory argument built from fundamentals usually has multiple implications, at least some of which are testable. The research effort is not complete until empirical tests have shown that implications drawn from the argument are consistent with reality" (p. 38).

Here Geddes thinks that rational choice approaches are particularly powerful because they readily conform to the underlying logic of statistical inference.[3] Such approaches are easily able to be generalized and subject to clear tests (that is, easy to falsify).[4] Thus, the fact that rational choice approaches abstract from "the specifics of particular cases" and deduce hypotheses from a precise model makes them a powerful tool in theory building (p. 206).

Coming from a different angle, the Elmans' *Bridges and Boundaries* also confirms much of the conventional wisdom about what makes political scientists scientific. The essays in this collection agree that the biggest difference between historians and political scientists is political science's self-conscious use of scientific methodology.[5] The first section explicitly takes up these issues of philosophical inquiry. The second section turns to case studies by historians and political scientists, illuminating different approaches to the same cases. Jack Levy's chapter on "Explaining Events and Developing Theories" suggests that the difference between historians and political scientists turns on the difference between the "logic of discovery" and the "logic of confirmation." The former is important in constructing theory, but political science is preoccupied with empirically validating theory (pp. 79–81). Following this, the Elmans suggest that political scientists are nomothetic, that is, interested in discovering generalizable theories, whereas historians tend to be idiographic, that is concerned with particular and discreet events (p. 13). Political scientists are not interested in a *case,* but in *cases* that enable them to test theories, discover the links between variables, and rule out competing theories of explanation.

But even with such larger goals in mind, most authors in *Bridges and Boundaries* emphasize the importance of beginning with good questions, which themselves begin with a particular case or puzzle (Lebow, p. 113). In this way, the qualitative study of cases is central to developing theory, which is not the result of a single test, but rather part of a larger research program that seeks to advance our theoretical understanding (Elman and Elman, pp. 12–20). To this end, Andrew Bennett and Alexander George in "Case Studies and Process Tracing in History and Political Science" and the historian John Lewis Gaddis in "In Defense of Particular Generalization" argue that contingent generalizations drawn from a rich understanding of cases, rather than broad theoretical generalizations, may be more consistent with the scientific aspirations of the discipline. Thus these qualitatively oriented political scientists and historians refuse to relinquish the mantel of science to scholarship that tests hypotheses based upon statistical inference alone (pp. 138, 320).

Tradeoffs and research design

Situating the testing of theory as only one aspect of the scientific enterprise, Gerring, Collier and Brady, and Ragin in their three books insist upon the tradeoffs inherent in any research design. In carrying out research we must balance what we are doing, recognizing that not all methodological goals are equal. Most importantly, before testing a hypothesis we want to be clear about concepts and theories. After all, concepts and theories are not prefabricated things to be randomly picked up by the political scientist eager to test this or that hypothesis, but the very building blocks of social science—they are the basis of good causal inference. Thus just where we cut into the process—whether we start with propositions, cases, or concepts—depends on the particular research in which we are engaged. "There is no obvious point of entry for a work of social science, no easy way to navigate between what we know (or think we know) and what we would like to know 'out there' in the empirical world. One might begin with a hunch, a question, a clearly formed theory, or an area of interest" (Gerring, p. 22).

John Gerring's *Social Science Methodology: A Criterial Framework* begins from this perspective. Gerring offers a

https://doi.org/10.1017/S1537592705050486 Published online by Cambridge University Press

"criterial framework" as a unifying methodological foundation for social science. This framework breaks social science into three tasks: concept formation, proposition formation, and research design. Concepts address what we are talking about, propositions are formulations about the world, and research design is how we will demonstrate a given proposition. These tasks are interdependent: alterations in one will necessarily result in alterations in another.[6] Finding the proper balance among these tasks depends upon the research at hand.

Each task involves different *criteria.* If our task is concept formation, the *criteria* of a good concept include the concept's coherence, operationalization, contextual range, parsimony, analytical utility and so forth (p. 40). Gerring gives us a set of criteria for proposition formation and research design as well. For example, the criteria for a good proposition include its accuracy, precision, breadth, depth, and so on (p. 91). In forming concepts, positing propositions, or framing research design, we must weigh competing criteria. Making our proposition broader will necessarily mean that it is less accurate and less deep. This requires us to engage in tradeoffs among criteria, as well as between these different tasks. Just where we cut into the hermeneutical circle will depend on our research task. If we are working with well-formed concepts and theories, we might proceed to proposition formation and to testing our propositions with a particular research design. But for Gerring, "the hard work of social science begins when a scholar prioritizes tasks and criteria. It is this prioritization that defines an approach to a given subject. Consequently, the process of putting together concepts, propositions, and research design involves circularities and tradeoffs (among tasks and criteria); it is not a linear follow-the-rules procedure" (p. 23).[7]

David Collier and Henry Brady's *Rethinking Social Inquiry: Diverse Tools, Shared Standards* shares Gerring's insistence on the tradeoffs inherent in social science inquiry. This is particularly true, Collier and Brady argue, in developing theory. Indeed, the preoccupation in mainstream quantitative methods with deducing a hypothesis and *testing* it against new cases may well inhibit theory building and development; there is an important tradeoff between theoretical innovation and the rigorous testing of a theory (p. 224). Making these necessary tradeoffs is the first step of any research design. This move situates research design as part of an interactive process between theory and evidence, rather than treating *testing* as a single-shot game. Thus, for example, Collier and Brady are skeptical of the insistence upon a *determinate* research design (pp. 236–38); given the many problems of causal inference, no observational research design is likely to be truly *determinate* (p. 237). It is not simply that research questions will always be open to further investigation. More importantly, the insistence upon determinate research design is likely to have negative consequences on the quality of theory—scholars are likely to redesign theory in accord with its *testability* rather than looking for ways to get at "theory that scholars actually care about" (p. 238). *Testing* theory in accord with a determinate research design calls for a larger N in accord with the imperatives of statistical testing. This move, however, creates problems of causal inference and neglects fruitful exchanges between theory and evidence. Given this tradeoff, Collier and Brady call for research designs that yield *interpretable* findings that "can plausibly be defended" (p. 238). This can be done along many lines, including increasing the N in some circumstances. But it may also be the result of "a particularly revealing comparative design, a rich knowledge of cases and context," or "an insightful theoretical model" (p. 238). Recognizing these tradeoffs and connecting them to *interpretable* research design recognizes "multiple sources of inferential leverage" and may lead to better causal inference.

Charles Ragin pushes this line of thought in *Fuzzy-Set Social Science.* The first section critiques mainstream quantitative approaches in social science, which are contrasted with case-oriented research. As he did in *The Comparative Method,* Ragin takes aim in particular at the assumption in statistical inference of unit homogeneity and additive causation. These, Ragin suggests, are dubious assumptions that often mask crucial differences between cases and therefore pose problems for causal inference. While he does not explicitly invoke the language of tradeoffs, Ragin emphasizes the difference between research designs aimed at testing theory and those aimed at building, developing, and refining theory. Ragin sees the latter as involving an iterative process between theory and case construction that we cannot get from testing theory. In the second half of the book, Ragin explains fuzzy-set methods in more detail. They represent his latest innovation in constructing categories to get at empirical evidence (data sets) that is too often taken for granted. Fuzzy sets begin with the cases themselves, illustrating how data are constructed by the researcher. This approach sets up a dialogue between ideas and evidence, where concepts are refined and theory developed based on interaction with the data. Ragin argues that such dialogue leads to more reliable causal inference because it does not rest on the dubious assumptions of mainstream variable-driven quantitative methods. In fact, Ragin's book is perhaps most valuable as an extended discussion of the logic of inference.

These books suggest that the distinction between *confirmatory* research and *exploratory* research is not as stark as it appears; that both are part of the overall scientific method (Gerring, p. 23). They also call on scholars to recognize the inherent tradeoffs in any social science research. As Gerring argues, one should justify one's approach based on the kind of question one is asking. This is more than a commonsense piece of advice, given that much of political science is preoccupied by the confirmatory approach, which neglects the crucial stages of concept formation and refinement, as well as the development and exploration of theory. Such *exploratory* research is an essential part of the scientific process, as it provides the basis for building concepts and theories.

Sharp disagreements remain even among these scholars who seek a unified language and a unified set of standards within political science. But there is important common ground: all of these books pay far more careful attention to concept formation, theory development, and causal complexity than have past works. This is true even when the larger goal remains testing theory from a statistical perspective (as in Geddes). In general, qualitative research attends to these concerns far more carefully than quantitative research, which often has treated them as prescientific. However, it is not simply that qualitative scholars have been more attuned to such concerns. Concept formation, theory building, and research design involve, at root, qualitative judgments. The scientific enterprise may best be seen as a framework that necessarily entails all of these tasks. How we situate our research within this framework depends upon qualitative judgments about where we are in the framework and what we are trying to do. These choices should be justified, but they cannot be justified in quantitative terms. As I will elaborate below, quantitative tools are being developed to further such efforts; yet, in an interesting twist, these tools are complementing, even imitating the logic of qualitative analysis (Collier and Brady, pp. 12–13). In the end, good theory—not simply method—is what makes for good science.

Concept Formation and Refinement: The Foundation of Social Science

In *Social Science Methodology,* Gerring suggests that "concept formation concerns the most basic questions of social science: *What are we talking about?*" (p. 35). We cannot posit hypotheses, let alone have a consensus on whether hypotheses and theories are valid, if we do not know what we are talking about. Too often, new definitions are proffered with little regard for already existing ones or for how such concepts fit into the world more generally. The result is that scholars speak past one another, research does not cumulate, or ideas simply remain sloppy. Given this, Gerring notes, there is a temptation just to get on with it. This will not do: a "blithely empirical approach to social science" only adds to the trouble (p. 38).

Before we can elaborate our theoretical arguments, we need to conceptualize what we are talking about. This begins from our understanding of the world, and it is an essentially qualitative project. Indeed, the first step in concept formation is picking words to describe what we are talking about. Even if we go with accepted definitions, we are relying on past qualitative judgments. Assigning cases to our conceptual definition also requires qualitative judgments. As we examine cases, compare and contrast them, we refine their boundaries and gain important conceptual distinctions. This will give us both more homogeneous cases for proper comparison and a sharper conceptual understanding of what we are talking about (Ragin 2004, pp. 125–28). This qualitative work is foundational to quantitative or statistical tests of theory. It requires that we specify concrete criteria by which to measure concepts, so that our measurements will allow for the assignment of cases to particular categories based upon the operationalized concept (Ragin 2004, pp. 144–45). This is all central to carrying out rigorous testing of theory that other researchers can replicate (Geddes, pp. 146–47; Collier and Brady, p. 209). Even when scholars take off-the-shelf indicators to measure a concept, or apply a well-defined concept to new cases, such judgments are unavoidably qualitative. In fact, relying on off-the-shelf indicators and extending concepts to a greater number of cases so they may be properly tested raises concerns about conceptual stretching and measurement validity. Picking up concepts and applying them to new empirical situations may lead to invalid causal inference (Collier and Brady, pp. 202–3). Collier and Brady note that statistical tools are being developed to help scholars get at these issues from a quantitative perspective. But when to use such tools, given the tradeoffs inherent in any research, itself depends on a qualitative judgment (pp. 204–9).

Beyond causal inference, Collier and Brady, along with several of their coauthors (see especially the chapters by Gerardo Munck and Ragin), view the refinement of concepts as an end in itself; this effort is not simply a precursor to testing theory (although it is also that). Classifying our knowledge, creating typologies and categories, is as an important a contribution to science as causal inference (p. 203). As Brady argues in a chapter in *Rethinking Social Inquiry,* the preoccupation with causation in *testing* theory rejects the notion that classification, categorization, and typologies have any explanatory power (pp. 53–67). This is unwarranted. Take Max Weber's division of authority into three essential types: traditional, legal-rational, and charismatic. It orders the world for us, beginning with what is out there, which forms the basis of our theorizing about it. Weber's approach is inductive and descriptive, but, as Gerring argues, surely scientific (Gerring, pp. 118–24).[8]

This is because we cannot speak of questions of fact "without getting caught up in the language to describe such questions" (p. 38). We cannot, for example, say that democracy requires X, without being clear about *democracy* itself. And for Gerring, we cannot avoid abstract concepts like democracy and ideology. A social science that avoids such high-order concepts would lack the power to generalize and thus be reduced to "a series of disconnected facts and microtheories" (Gerring, p. 38). These higher-order concepts not only let us negotiate general propositions; they provide the context in which we situate

https://doi.org/10.1017/S1537592705050486 Published online by Cambridge University Press

lower-order concepts that are more readily observable in empirical terms. It would be odd, would it not, to study *voters* and *political participation* if we could not situate these lower-order concepts within *democracy*? In this sense, "concept formation in the social sciences may be understood as an attempt to mediate between the world of language . . . and the world of things" (Gerring, p. 37). We cannot speak of the stuff of the world without concepts, but our understanding of this "stuff out there" will help us refine and alter our concepts.

While political scientists speak of the need for agreed-upon scientific standards—the need for generalizability and a common language—we often fail in this goal by treating concept formation in haphazard manner. Here Gerring's development of a *min-max* definition for concepts is useful in providing a foundation for a common language, giving us the ability to generalize across a wide range of issues. Gerring breaks concept formation into three steps: "sampling usages, typologizing attributes, and the construction of minimal and ideal definitions" (p. 71). The idea is to construct a conceptual definition that fits within the field of existing usage, avoiding highly idiosyncratic views, but also organizing the conceptual definition in such a way that allows for general application. To do this, Gerring urges constructing a min-max definition. It begins with the minimal features of a concept—its core—giving us a capacious definition by "casting the semantic net widely" (p. 78). From here, we construct an *ideal type* definition that captures the concept in its purest form. This is, of course, a purely logical construct, far less likely to be found in the empirical world. These two poles—the minimum and maximum of a concept—allow us to broaden and narrow the term. The minimal definition (with a smaller intension) captures more "out there" (the extension), whereas a maximum definition (with a much larger intension) narrows the cases to which it will apply (the extension). The result is a definition that "outlines the parameters of a concept," giving us a "frame within which all contextual definitions should fall." This allows us to move from a general definition to a more contextual definition while maintaining a common ground. It also enables our conceptual definition to travel through time and space.

Theory Building: Development and Exploration

The relationship between theory and evidence is often seen as the key distinction between science and non-science in the study of politics. And while these authors offer a variety of advice on selecting cases to bolster theory, they share an important point of agreement. Increasing the number of cases or observations is not nearly as important as carefully selecting cases, which often depends upon substantive knowledge of the cases. Even if the intent is to test a hypothesis against evidence, a small number of cases that maximize the researcher's leverage may be far superior to increasing the number of cases, which risks creating problems of conceptual stretching and measurement validity.

It is almost a mantra among social scientists that a true test of one's theory must be separate from any evidence used in generating the theory. Geddes's *Paradigms and Sand Castles* extols this conventional wisdom to great effect, urging that scholarship be carried out with more "sensitivity to the norms of science" (p. 144). To more readily share data, test arguments, and replicate tests, case-study evidence needs to be more carefully operationalized and measured. This does not just mean specifying the domain of cases to which an argument applies. We must carefully specify the criteria of measurement and the classification and categorization of cases. We must then stick to these publicly articulated criteria across cases. This, Geddes argues, can be arduous, but it is well rewarded: "struggling over the conceptual issues raised by trying to measure the same causal factors in the same way across cases often deepens the analyst's understanding of the argument as well as the cases" (p. 147). Thus, even while endeavoring to test cases, Geddes illustrates the important interplay between theory and evidence. Geddes warns against assuming an "accumulation of theory" that leads scholars to neglect the importance of building theory.[9] She also ventures that the accumulation of theoretical knowledge will only occur if scholars are far more careful in using the "inadequate and fuzzy evidence" at hand to *test* theory. But what happens when confrontations between theory and evidence do not yield straightforward results?

Most of the authors here suggest that this is, in fact, the norm. This leads them to suggest a continual movement between theory and evidence that clarifies and refines theory. Theories from which we derive propositions are constructed based on our interaction with the empirical world. This interaction is central to the scientific enterprise—not a prescientific aspect of it. Furthermore, building theory is done against the backdrop of what we already know or think we know. Thus Timothy McKeown invokes the hermeneutic maxim: "no knowledge without foreknowledge," arguing that tests of theory are not "pointlike observations" but "iterated games" situated against existing theory.[10] As McKeown notes:

> If the extent of one's knowledge about political science were the tables of contents of most research methods books, one would conclude that the fundamental intellectual problem facing the discipline must be a huge backlog of attractive, highly developed theories that stand in need of testing. That the opposite is more nearly the case in the study of international relations is brought home to me every time I am forced to read yet another attempt to "test" realism against liberalism (McKeown, 24–25; see also Ragin 2004, 127).

Substitute your field of choice and you are likely to be confronted with the same problem.

Learning from our cases in the process of building theory is most closely linked with qualitative tools and scholarship, but the distinction between theory and evidence also touches upon an important debate between *quantitative* scholars. It is a central part of Collier and Brady's statistical critique of mainstream quantitative methods, which contrasts Bayesians with frequentists. Mainstream regression analysis, which provides Geddes's underlying framework, begins from a frequentist approach that assumes we know nothing about the world. A model is elaborated, hypotheses specified and tested against the data (cases).[11] The data are thus used to "construct the world." In this approach a "lack of context-specific knowledge means that the researcher cannot call on information from outside the sample being analyzed to supplement the information gleaned from a statistical analyses" (McKeown, p. 148). By contrast, in a Bayesian approach, prior knowledge about the world is assumed to be deeply important, and the data are important only to the extent that they modify our prior assumptions or theories. Thus cases (data) are weighed against prior understandings and not a blank slate (null hypothesis). Drawing on this logic, Collier and Brady illustrate how statistical theory incorporates such notions, imitating the logic of qualitative research in carefully selecting cases against existing theory (pp. 233–35). For example, when McKeown refers to qualitative researchers as *folk Bayesians* he is drawing attention to the fact that qualitative researchers usually situate their specific research within a larger research framework—that is, against conventional theoretical understandings—and select their cases with this in mind.[12]

From a Bayesian perspective, the inductive refinement of theory is widely recognized as an essential part of research; it must be distinguished from data mining, where a hypothesis is deductively posited and data are selectively gathered to support it. In fact, Collier, Brady, and Seawright suggest that the indiscriminate injunction against inductive procedures "[and] even worse, the misleading pretense that they are not utilized" in statistical theory, is a much larger problem in social science than the legitimate use of induction (see also Munck pp. 119–20, Gerring, pp. 240–43). Such methods are widespread in more advanced statistical approaches (Collier, Brady, and Seawright, pp. 232–44). Specification searches, for example, rest on the logic of induction and are seen as a great innovation in statistical analysis. Moreover, rejecting these tools, particularly in qualitative research, may lead to boring political science (see also Gerring, pp. 240–43). The methodological norms that Geddes advocates, for example, may make it harder to develop interesting theory that scholars actually care about.[13]

Case studies, typologies, and descriptive work are invaluable in this effort (Gerring, p. 122). As we refine cases, we refine theories—not simply to fit the evidence, but to get a better understanding of cases and concepts, allowing us to construct better theory (Ragin 2004, p. 127). We begin with preexisting theories that will lead us to select cases based on substantive knowledge of the field rather than random selection. "[W]hat guides research is not logic but craftsmanship, and the craft in question is implicitly far more substantively rich than that of 'social scientist without portfolio'" (Mckeown, p. 148; see also Geddes, p. 31). It is too bad that McKeown downplays logic here. For as Ragin and others argue, the best case studies most certainly follow the rules of logic—a point McKeown himself acknowledges when referring to qualitative researchers as *folk Bayesians.*

Consider Randall Schweller's chapter in *Bridges and Boundaries.* Schweller uses World War II as a test case for a "systems theory of international relations" that focuses on the distinct features of a tripolar system (pp. 196–97). He selects World War II because "it contradicts many of the key behavioral predictions of realist balance of power theory" (p. 182). Thus Schweller selects his case based on his substantive knowledge of it and against existing theory. Recognizing that many states did not balance against threats between World Wars I and II in accord with realist balance-of-power theory, Schweller puts forward new hypotheses that suggest a modification of existing theory. As he explains, "World War II is an important case for international relations theory because it suggests ways to revise, reformulate, and amend realist balance of power theory 'for the purposes of (1) better explanations and more determinate predications, (2) refinement and clarification of the research program's theoretical concepts, and (3) extension of the research program to cover new issue areas" (p. 182).

This movement between theory and evidence even in the course of testing theory leads Ragin to suggest that a case (or cases) is best thought of as a working hypothesis to be refined in the course of the research. We begin against the backdrop of prior theory. In the course of the research, however, the boundary of the set of relevant cases "is shifted and clarified," and "usually this sifting of cases is carried on in conjunction with concept formation and elaboration" (Ragin 2000, p. 58). Schweller, for example, begins with World War II as a breakdown of the international system, but one that did not lead to balancing among great powers. His explanation turns on an emergent tripolar system, and he ends up classifying World War II as a case of tripolarity. This is all part of the refinement, elaboration, and development of theory. In short, while statistical (or variable-oriented) research turns to a predetermined set of cases to test a hypothesis, case-oriented research begins with a case (or cases), "treating them as singular whole entities purposefully selected, not as homogeneous observations drawn from a random pool of equally plausible selections" (Ragin 2004, p. 125).

https://doi.org/10.1017/S1537592705050486 Published online by Cambridge University Press

Political science research too often "assumes a pre-existing population of relevant observations, embracing both positive and negative cases, and thus ignores a central practical concern of qualitative analysis—the constitution of cases" (Ragin 2004, p. 127). Cases are not prefabricated. They must be constructed and refined before theory can be tested: "in case oriented research, the bulk of the research effort is often directed toward constituting 'the cases' in the investigation and sharpening the concepts appropriate for the cases selected" (Ragin 2004, p. 127). In this sense, the statistical template is parasitic upon qualitative research that constitutes cases, or it relies on prefabricated data sets drawn from "out there" (see Collier, Brady, and Seawright, pp. 252–64), which themselves rest upon qualitative judgments at some level.

It is this random sampling of data to *test* theory that leads to standard injunctions about case selection proffered from a mainstream quantitative template: never select on the dependent variable; one cannot draw causal inferences from a single case study; and the like. Such advice may well apply to standard regression analysis. But against the backdrop of prior knowledge and theory, a single *critical* case can confirm or refute existing theory. Selection on the dependent variable may be useful in testing for necessary and sufficient causation, examining causal mechanisms within a particular case, or refining concepts and cases themselves (Collier and Brady, McKeown, Ragin). The question in all of these instances is: What knowledge are we trying to gain? Where are we in the research process? Seeing theory building as an iterated game (McKeown), as a dialogue between ideas and evidence (Ragin), as situated in an overall research program (Elmans), or as about tradeoffs (Gerring and Collier and Brady), these authors offer a more capacious view of the scientific process. This understanding of the scientific process, moreover, will lead to better causal inference, particularly in the realm of causal complexity

Getting a Handle on Causal Complexity

Qualitatively oriented political scientists have turned to complex causality as more representative of the social world than conventional statistical inference. These scholars are often engaged in macro inquiry, large comparisons, situating politics within time, or seeking to get at causal mechanisms within a case. Such qualitative studies have sought to maintain social science standards rather than simply abandoning the enterprise and turning to historical description. Recognizing that conventional statistical inference does not apply to some of the most important questions within political science, these authors illustrate how qualitative tools are invaluable in getting at causal complexity. Moreover, the most sophisticated quantitative scholars are following this logic.

Standard statistical views seek to discover probable causal relationships: if X occurs, then it is likely that Y will too. There may be other causes of Y, and X may even act with other causes, but X precedes Y more than most other causes. This view rests on two important assumptions, that these variables all act in an additive fashion (so that X acts the same with A as it does with B), and that this is a linear relationship. How X acts with A is the same at T1 as at T2. It is this view of causation that leads to Geddes's advice about breaking theories down into smaller hypotheses, selecting cases, and using evidence to test theory. Most importantly, it is this view of causation that leads to the general insistence that increasing the number of cases (observations) tested against a hypothesis is the primary means of causal inference.

Attempts to isolate and measure a single variable—assuming uniformity across cases—may be justified in some circumstances, but they are unlikely to capture causal complexity.[14] And as Ragin argues, causation is often the result of conjunctions and interactions between variables. That is, X causes Y in conjunction with A, B, and C. Or Z causes Y when A, B, and C are absent. Thus Ragin rejects the notion that a causal variable is likely to act the same across a wide range of cases. Rather, X is likely to act in conjunction with other variables or conditions. Attempts to assess X's impact across a wide range of cases in isolation (a standard regression model) is likely to miss the significance of X if it works in conjunction with other variables. Getting at complex causality begins with the qualitative construction of cases, defining both positive and negative cases and exploring the causal factors that produce these outcomes.

Qualitative methods are well developed for examining critical-juncture, path-dependent, and nonlinear causes, which are increasingly at the heart of comparative historical analysis, American political development, and international relations.[15] Geddes, for example, recognizes the importance of causal complexity with critical juncture and path-dependent arguments. But these forms of causation elude her quantitative template. She argues that "the big argument should be broken at its branching points, and hypotheses that purport to explain choices at different nodes should be tested on additional cases that fit appropriate initial conditions. Efforts should be made to see if hypothesized legacies also occur in these other cases" (p. 14; see also p. 23). Geddes's advice to break down big questions into smaller questions is driven by her view of causal inference: she wants to increase the number of observations against any particular hypothesis because she is working from a mainstream quantitative template. This seems a common-sense piece of advice.

But as McKeown and others argue, such a move represents the limits of a statistical world view. Moving away from big questions to those that can be tested within a statistical framework will not capture complex causality

https://doi.org/10.1017/S1537592705050486 Published online by Cambridge University Press

(McKeown, Ragin, Gerring, Collier and Brady).[16] If the world is characterized by complex multi-causality, isolating individual variables and testing them across time will not capture how they interact; indeed, it is inconsistent with the ontology of complex causality.[17]

This is evident in what Robert Jervis has called system effects.[18] How each piece of the puzzle is situated in a particular system matters greatly—far more than any individual piece. "The importance of interactions means that the common habit of seeing the impact of variables as additive can be misleading: the effect of A and B together is very different from the effect that A would have in isolation and added to the impact of B alone."[19] In *Politics in Time,* Paul Pierson focuses on the temporal context that variables act within, arguing that it matters *when* something happens. Different outcomes may be the result of the conjunction of forces at a particular time, giving us *critical junctures.* How X, Y, and Z interact at T1 may be very different from how they interact at T2. A specific conjunction combined with other forces may lead to divergent outcomes, giving us branching causes. X combined with AB may lead to one pathway, while X combined with DE may lead to another. Or we may get *path-dependent* causes. The fact that X occurred at T1 makes it much more probable that Y and not Z will occur afterward. Wresting these variables from their temporal context means that we will miss these interactions. Moreover, a *snapshot* approach that focuses on micro-causes may fail to capture slow-building causal processes where X triggers A, B, and C, which results in Y. This process will only become evident over an extended period of time. In this sense, standard variable-centered analysis is the equivalent of a snapshot of the political world. Situating variables in their temporal context—getting a handle on the process and sequence of events—provides a moving picture rather than a snapshot.[20]

Process tracing, as detailed by George and Bennett in *Bridges and Boundaries,* is another powerful tool that helps illustrate complex causal mechanisms. By tracing the sequence of events, process tracing can explain *how* X causes Y or, far more likely, how Y is brought about by various combinations of Xs. Process tracing is rooted in case studies and midlevel theory—often engaging in typologies—that identify contingency and variation, balancing rich investigation with more expansive theorizing. Collier and Brady similarly suggest that thick analysis and causal process observations may be more fruitful in getting at causal mechanisms than thin analysis and data set observations (pp. 248–49, 252–64). As with other qualitative tools, process tracing can inductively get at causal mechanisms, compare causal pathways, and even permit causal inference on the basis of a few cases. Probabilistic views of causation, while making a nod to the value of process tracing under certain conditions, suggest that we do not need to get at the specifics of how X causes Y; it is enough to see that a general relationship holds. This leads KKV, for example, to suggest that causal effects are by definition prior to causal mechanisms.[21] Such thinking is especially evident in the search for covering laws and general theories that discount tracing the links between each individual X and Y. Yet, even if such covering laws are true in probabilistic terms, they often operate at such a level of abstraction that they are not particularly insightful.

Consider the presidential election of 2000. Almost all political science theories predicted that a vice president running on peace and prosperity with a sound economy would win in a landslide.[22] And, indeed, as a probabilistic law this would almost certainly hold true.[23] It doesn't tell us much, though, does it? In the idiom of statistics, outliers and anomalous findings may well be the most interesting cases. Thus we are far more interested in the particular circumstances of the 2000 election: what variables were at work, in conjunction with what circumstances, to bring about this result? To get at these more interesting questions—why did the 2000 presidential election result in a perfect tie?—we must recognize particular circumstances and combinations of variables, not broad covering laws.

This is true even when a model is more accurate in predicting results.[24] Even if we predict results based upon the scores of a wide variety of variables we have not explained *why* this happened. Such *data set* observations may well predict that the scores of a set of variables—for example, incumbency, duration of the party in power, peace, economic growth, good news, inflation—is likely to result in, say, a small margin of victory for a sitting vice-president.[25] But this does not explain how these variables work to bring about a given result. This is, however, what causal process observations attempt to do. Take an illuminating example. In the middle years of the Clinton administration, Stephen Skowronek ventured that President Clinton, given the conjunction of forces that aligned in the Republican takeover of Congress in 1994 and his reelection, was a likely candidate for impeachment.[26] Skowronek was theorizing based on similar past conjunctions—a typology of presidential regimes—that, to put it bluntly, looked preposterous in 1997. Of course by 1998, such contingent theorizing looked extraordinarily impressive, far more so than the statistical models that generally predicted a smashing Gore victory in 2000.[27]

The statistical view of causation is being challenged from a quantitative angle as well. Collier and Brady, for example, note that a probabilistic view of causation is ill applied when we are dealing with deterministic views of causation (where X is either a necessary or sufficient cause). Yet deterministic views of causation are significant in political science research. If one is working with a deterministic view of causation, there is not a problem with selecting on the dependent variable or using a *critical case* to falsify a theory. If X is deemed necessary to bring about Y, the cases that provide the test for the hypotheses will necessarily include Y—giving us no variation on the dependent variable, but allowing for causal inference. Or if we find X but no Y in a *crucial* case, that is grounds for rejecting the hypothesis that X (as necessary and sufficient) causes Y, for we have X but not Y. As Geddes argues, "the analyst should understand what can and cannot be accomplished with cases selected" (p. 90).

Thus Collier and Brady insist that the contributions of traditional qualitative methods and qualitative critiques of statistical methods can be justified from the perspective of statistical theory. This terminology can be confusing and may belie the notion that we are moving closer to a shared language in political science. Collier and Brady, for example, argue that advanced statistical theory illustrates that statistical inference (Geddes) cannot provide an adequate foundation for political science. What Geddes and most of the authors here call statistics Collier and Brady refer to as mainstream quantitative analysis (one branch of statistics based largely on regression analysis). Collier and Brady make a powerful case, noting that attempts to impose a mainstream quantitative template upon qualitative research (for example, Geddes and KKV) neglect crucial insights and developments from more *quantitative* approaches.

Collier, Brady, and Seawright note how statistical tools are being developed to better address these problems of causal inference and causal complexity—Bayesian statistical analysis being a leading example. Similarly, Ragin has used tools like Boolean algebra which compare different combinations of variables to get at similar outcomes in multiple-conjunctural causation. Fuzzy-set methods are another tool that illuminates necessary and sufficient causation and combinations of variables, using probabilistic criteria. These methods often employ statistics and quantitative analysis, so they do not easily fit the quantitative-qualitative divide. Indeed, Collier, Brady, and Seawright insist that in assessing causation and "evaluating rival explanations, the most fundamental divide in methodology is neither between qualitative and quantitative approaches, nor between a small N and a large N. Rather, it is between experimental and observational data" (p. 230). In this, they offer a rather damning indictment of attempts to impose a mainstream quantitative template upon all political science research. They take aim, in particular, at KKV (and, by implication, Geddes). For while KKV purport to offer advice to researchers engaged in observational studies—political science is the business of *observing* the political world—their view of causation "takes as a point of departure an experimental model" (p. 232). Thus if Collier and Brady are correct, those engaged in small N qualitative studies cannot solve the problem of causal inference by seeking to imitate large N quantitative studies, as these problems are inherent in large N studies as well (p. 233).

This is what leads Collier and Brady to call for *interpretable* research design. Can the findings and inferences of a particular research design plausibly be defended? Such a defense must be justified in qualitative terms; as the failed quantitative template reveals, there is no methodological Rosetta stone that will solve all our problems. Yet it is fitting that much of the advice that these books proffer is of the nuts and bolts variety. They seek to illuminate the different tools at our disposal in seeking to understand politics. But which tools we use, when, and how we use them, depend on the judgments of each scholar given the research he or she is engaged in. In the end, there is nothing to do but justify our research and defend the results. Good scholarship is based, if I may repair to the words of an earlier political scientist, upon "reflection and choice."[28]

Conclusion

This review essay has argued that these five books give us a better understanding of the scientific process, especially concept formation and refinement, theory building and exploration, and issues of causal complexity. In taking up these issues in great detail, these works rescue important aspects of the vocation of political science that are often treated as prescientific. In restoring these features to their proper place in the scientific enterprise, these works challenge the dominant view of science within the discipline. That is, political science's quest for a common vocabulary and set of standards cannot be found in statistical inference; if political science is to have a unified methodology, that methodology has qualitative foundations. This conclusion does not simply replicate the quantitative-qualitative divide. On the contrary, as I have attempted to show, many of the concerns associated with qualitative methods increasingly find expression in quantitative forms. Conversely, concept formation and theory building are just as important to quantitative researchers as to qualitative researchers. And while testing theory against evidence is one dimension of this process, it is not the whole of (political) science. Nor is it the truly *scientific* part.

By situating the empirical testing of theory as one part of an ongoing process, these works give us a more thorough, capacious, and practical view of science itself. What questions does our research seek to illuminate? The key is to recognize where we are in a given research program and begin with the tasks at hand. Quantitative tools may help us in these tasks, but forming concepts and building theories rely upon qualitative judgments. If we think of the scientific process as one that inherently involves tradeoffs, the choices we make need to be justified. Indeed, just where we cut into the hermeneutical circle is a qualitative judgment.

The exhortation to a more systematic and rigorous view of science has led political scientists to turn to quantification, rational choice theory, and ever more sophisticated

https://doi.org/10.1017/S1537592705050486 Published online by Cambridge University Press

mathematical modeling. Thus the dominant view of science within political science has been drawn from physics, chemistry, and, more and more, math.[29] Yet this exhortation to be more scientific has now led to a rebellion against this narrow conception of science; these books illuminate how science itself has qualitative foundations. What was once effect has now become cause, perhaps revealing that political science itself moves in a complex and nonlinear way. Why should we expect the empirical world out there to be any different?

Notes

1 King, Keohane, and Verba 1994, 4.

2 Albeit in a "softened" form.

3 Thus Geddes' rational choice approach rests upon the conventional quantitative template.

4 This point is heavily debated as many criticize rational choice theories as tautological. See Green and Shapiro 1996.

5 Though see the excellent chapter by John Lewis Gaddis, who doubts that political scientists are in fact more scientific. Elman and Elman 2003a, 301–26.

6 Alternatively, these may be viewed as hierarchical; propositions are derivative of concepts, while research designs, going farther down the scale, are derivative of propositions. But even if this is philosophically so (as with questions of ontology and epistemology), it is not necessarily how we delve into our research; nor does it capture the triangular and interrelated relationship among these tasks.

7 See also Sartori 1984.

8 This also has the virtue of returning political theory to the fold of political science. Indeed, it takes seriously the claim that Aristotle is the first political *scientist.*

9 While KKV and Geddes insist on a Popperian view of "testing" theory, they miss the fact that Karl Popper's scientific method began with an inductive core of observation, from which theory and knowledge were built.

10 To Geddes and some of the authors in *Bridges and Boundaries,* theory building is really the fruit of testing or confirming theory. Cases are important in constructing theory and in providing the evidence against which a theory may be tested. But these are treated as prescientific or subordinate to the scientific goal of testing theory.

11 Geddes does note the importance of Bayesian approaches, 114–17.

12 It might be more accurate to describe Bayesians as qualitative-quantifiers, as qualitative scholars were utilizing such approaches long before Bayesian analysis.

13 See especially Rogowski's (1995) exchange with KKV in the APSR symposium, which is extended in Rogowski's chapter in Rethinking Social Inquiry. See also the exchange between Gerring 2003, Skowronek 2003, and Smith 2003a.

14 These assumptions were warranted in the study of American electoral behavior, where they were first used to great effect.

15 Mahoney 2003; Skowronek and Orren 2004; Jervis 1998 and 2000.

16 See also Mahoney and Rueschemeyer 2003 and Shapiro 2004.

17 Hall 2003. At a deeper level, scholars like KKV want to insist on a transcendent methodology without theories of knowing or assumptions about what the world is composed of. But just as concepts are related to propositions and particular research designs, methodology is derivative of epistemology, which is derivative of ontology. One cannot insist on a methodology without having epistemological and ontological assumptions.

18 Jervis 1998.

19 Jervis 2000, 95.

20 Pierson 2000a and 2004.

21 This becomes an argument between positivism and realism in the philosophy of science, which is beyond the scope of this essay.

22 See the various post election ruminations and adjustments to models in *PS: Political Science & Politics* 34 (1), 1–44.

23 Though this would also miss the election of 1960.

24 See, Ray Fair, http://fairmodel.econ.yale.edu/vote2004/index2.htm, which came closer to predicting the 2000 election. Yet his prediction illustrates the relevance of context. He predicted Gore winning the *popular vote* by a small margin and doesn't take up the Electoral College (but presuming that a popular win nationwide will translate in an electoral victory). Yet, in an interesting way, there is no popular vote nationally. Yes, the total number of votes cast nationwide is counted and reported, but that is not how we actually vote as a nation. A point vividly brought home in the 2000 election. Rather than voting in a national plebiscite, voting is conducted on a state-by-state basis. These institutional rules not only influence how candidates campaign, but have a profound impact on the elections—particularly close ones. Thus one cannot assume in a close election that more votes nationwide will necessarily result in an Electoral College victory. Attending to this contextual issue may have led to a more accurate prediction.

25 This is what Ray Fair's model essentially predicted for the 2000 election. But Fair's methodology is more akin to Collier and Brady's discussion of

https://doi.org/10.1017/S1537592705050486 Published online by Cambridge University Press

statistical theory than to the standard quantitative template of KKV. Indeed, Fair's model seems a great example of Collier and Brady's insistence that advanced quantitative methods often use inductive procedures in terms of specification searches and the like (thus imitating traditional qualitative methods). Moreover, assigning weights to these different variables, particularly over time in refining the model based upon an interaction with the evidence, depends upon qualitative judgments. For Fair's discussion, see http://fairmodel.econ.yale.edu/RAYFAIR/PDF/2002DHTM.HTM.

26 Skowronek 1997.

27 The Fair model is, again, a general exception, although not necessarily very illuminating in explaining *why* this occurred.

28 See Hamilton, Madison, and Jay 1999, 1.

29 This is odd. It treats such established hard sciences as biology and geology as almost unscientific, as they do not follow the logic of physics and math. This is all the more peculiar in that mathematics, unlike biology, is not even concerned with the empirical world—the fundamental preoccupation of political science. The methodology on display in these sciences might be better suited to social scientists than are physics or chemistry.

References

Adcock, Robert and David Collier. 2001. Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review* 95 (3): 529–46.

Bennett, Andrew. 2003. A Lakatosian reading of Lakatos: What can we salvage from the hard core? In *Progress in international relations theory,* ed. Colin Elman and Mirhiam Fendius Elman, 455–94. Cambridge: MIT Press.

Bennett, Andrew, Aharon Barth, and Kenneth Rutherford. 2003. Do we preach what we practice? A survey of methods in political science journals and curricula. *PS: Political Science and Politics* 36 (3): 373–78.

Blyth, Mark. 2002a. *Great transformations: Economic ideas in the twentieth century.* New York: Cambridge University Press.

———. 2002b. Institutions and ideas. In *Theory and methods in political science,* ed. David Marsh and Gary Stoker, 293–310. New York: Palgrave Macmillan.

———. 2003. Structures do not come with an instruction sheet: Interests, ideas, and progress in political science. *Perspectives on Politics* 1 (4): 695–706.

Bridges, Amy. 2000. Path dependence, sequence, history, theory. *Studies in American Political Development* 14 (1): 109–12.

Caporaso, James A. 1995. Research design, falsification, and the qualitative-quantitative divide. *American Political Science Review* 89 (2): 457–60.

Ceaser, James W. 1990. *Liberal democracy and political science.* Baltimore: Johns Hopkins University Press.

Collier, David. 1995. Translating quantitative methods for qualitative researchers: The case of selection bias. *American Political Science Review* 89 (2): 461–66.

Eckstein, Harry. 1975. Case study and theory in political science. In *Handbook of political science,* vol. 7, ed. Fred I. Greenstein and Nelson W. Polsby, 79–138. Reading, MA: Addison-Wesley.

Elman, Colin, and Miriam Fendius Elman. 2003a. Lessons from Lakatos. In *Progress in international relations theory,* ed. Colin Elman and Miriam Fendius Elman, 21–68. Cambridge: MIT Press.

Elman, Colin and Miriam Fendius Elman, ed. 2003b. *Progress in international relations theory: Appraising the field.* Cambridge: MIT Press.

Gerring, John. 2003. APD from a methodological point of view. *Studies in American Political Development* 17 (1): 82–102.

Green, Donald P. and Ian Shapiro. 1996. *The pathologies of rational choice theory: A critique of applications in political science.* New Haven: Yale University Press.

Hall, Peter. 2003. Aligning ontology and methodology in comparative research. In *Comparative historical analysis in the social sciences,* ed. Mahoney and Rueschemeyer, 373–404. New York: Cambridge University Press.

Hamilton, Alexander, James Madison, and John Jay. 1999. *The Federalist papers.* New York: Mentor Books.

Jervis, Robert. 1998. *System effects: Complexity in political and social life.* Princeton: Princeton University Press.

———. 2000. Timing and interaction in politics: A comment on Pierson. *Studies in American Political Development* 14 (2): 93–100.

King, Gary. 1995. Replication, replication. *PS: Political Science and Politics* 28 (3): 443–99.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing social inquiry: Scientific inference in qualitative research.* Princeton: Princeton University Press.

———. 1995. The importance of research design in political science. *American Political Science Review* 89 (2): 475–81.

Laitin, David. D. 1995. Disciplining political science. *American Political Science Review* 89 (2): 454–56.

———. 2003. The Perestroikan challenge to social science. *Politics and Society* 31 (1): 163–85.

Lijphart, Arend. 1971. Comparative politics and the comparative method. *American Political Science Review* 65 (3): 682–93.

https://doi.org/10.1017/S1537592705050486 Published online by Cambridge University Press

Mahoney, James. 1999. Nominal, ordinal, and narrative appraisal in macrocausal analysis. *American Journal of Sociology* 104 (4): 1154–96.

———. 2003. Strategies of causal assessment in comparative historical analysis. In *Comparative historical analysis in the social sciences,* eds., James Mahoney and Dietrich Rueschemeyer, 337–72. New York: Cambridge University Press.

Mahoney, James, and Dietrich Rueschemeyer, eds. 2003. *Comparative historical analysis in the social sciences.* New York: Cambridge University Press.

Mansfield, Harvey, Jr. 1990. Social science versus the Constitution. In *Confronting the Constitution,* ed. Allan Bloom, 411–36. Washington, DC: American Enterprise Institute Press.

Pierson, Paul. 2000a. Not just what, but when: Timing and sequence in political processes. *Studies in American Political Development* 14 (2): 72–92.

———. 2000b. Increasing returns, path dependence and the study of politics. *American Political Science Review* 94 (2): 251–67.

———. 2003. Big, slow-moving, and . . . invisible. In *Comparative historical analysis in the social sciences,* ed., James Mahoney and Dietrich Rueschemeyer, 177–207. New York: Cambridge University Press.

———. 2004. *Politics in time: history, institutions, and social analysis.* Princeton: Princeton University Press.

Ragin, Charles C. 1987. *The comparative method: Moving beyond qualitative and quantitative strategies.* Berkeley: University of California Press.

———. 2004. Turning the tables: How case-oriented research challenges variable-oriented research. In *Rethinking social inquiry: Diverse tools, shared standards*, eds. David Collier and Henry Brady. Lanham: Rowman and Littlefield.

Rogowski, Ronald. 1995. The role of theory and anomaly in social-scientific inference. *American Political Science Review* 89 (2): 467–70.

Sartori, Giovanni. 1984. Guidelines for concept analysis. In *Social science concepts: A systematic analysis,* ed. Giovanni Sartori, 15–85. Beverly Hills, CA: Sage Publications.

Shapiro, Ian. 2004. Problems, methods, and theories in the study of politics, or: What's wrong with political science and what to do about it. In *Problems and methods in the study of politics,* eds. Ian Shapiro, Rogers M. Smith, and Tarek E. Masoud, 19–41. New York: Cambridge University Press.

Skocpol, Theda, and Margaret Somers. 1980. The uses of comparative history in macrosocial inquiry. *Comparative Studies in Society and History* 22 (2): 174–97.

Skowronek, Stephen. 1997. *The politics presidents make: Leadership from John Adams to George Bush.* Cambridge: Harvard University Press.

———. 2003. What's wrong with APD? *Studies in American Political Development* 17 (1): 107–10.

Skowronek, Stephen, and Karen Orren. 2004. *The search for American political development.* New York: Cambridge University Press.

Smith, Rogers M. 1996. Science, non-science, and politics. In *The historic turn in human sciences,* ed., Terrance J. McDonald, 119–59. Ann Arbor: University of Michigan Press.

———. 2003a. Substance and methods in APD research. *Studies in American Political Development* 17 (1): 111–15.

———. 2003b. Progress and poverty in political science. *PS: Political Science and Politics* 36 (3): 395–96.

Tarrow, Sidney. 1995. Bridging the quantitative-qualitative divide in political science. *American Political Science Review* 89 (2): 471–74.

Thelen, Kathleen. 2000. Timing and temporality in the analysis of institutional evolution and change. *Studies in American Political Development* 14 (1): 101–8.