Title
A Sea of Change in Political Methodology?

Journal
Qualitative and Multi-Method Research, 9(1)

Authors
Collier, David
Brady, Henry E
Seawright, Jason

Publication Date
2011-04-01

Peer reviewed

***Qualitative and Multi-Method Research*. Newsletter of the American Political Science Association Organized Section. 9, No. 1 (Spring 2011, forthcoming).**

A Sea Change in Political Methodology[1]

David Collier
University of California, Berkeley
dcollier@berkeley.edu

Henry E. Brady
University of California, Berkeley
hbrady@berkeley.edu

Jason Seawright
Northwestern University
j-seawright@northwestern.edu

Shifting debates on what constitutes "science" reveal competing claims about methodology.[2] Of course, in its origin the term "science" means "knowledge," and researchers obviously hold a wide spectrum of positions on how to produce viable knowledge. Within this spectrum, we compare two alternative meanings of science, advanced by scholars who seek to legitimate sharply contrasting views of qualitative methods. This comparison points to a sea change in political science methodology.[3]

[1] This article draws on the Introductions to Parts I and II of Brady and Collier, *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2nd edn. (Lanham, MD.: Rowman & Littlefield, 2010).
[2] Morgan (1996) provides a broad overview of rival views of science, encompassing the natural, biological, and social sciences.
[3] For our own work, we share David Freedman's (2010a) view of plurality in scientific methods, and we also recognize social versus natural science as partially different enterprises. Yet the two can and should strive for careful formulation of hypotheses, intersubjective agreement on the facts being analyzed, precise use of data, and good

One meaning of “science” is advanced in King, Keohane, and Verba’s (KKV) 1994 book, *Designing Social Inquiry*, which proposes a bold methodological agenda for scholars who work in the qualitative tradition. The book’s subtitle directly summarizes the agenda: “*scientific* inference in qualitative research” (italics added). To its credit, the book is explicit in its definition of science. It draws on what we and many others have viewed as a “quantitative template,” which serves as the foundation for the desired scientific form of qualitative methods. In KKV’s view, standard research procedures of qualitative analysis are routinely problematic, and ideas drawn from conventional quantitative methods are offered as guideposts to help qualitative researchers be scientific.

A starkly different position has been emerging, forcefully expressed by the statistician David A. Freedman (2010a). He reviews the central role of qualitative analysis in six major breakthroughs from the history of epidemiology—a field highly relevant to political science because it faces many of the same challenges of causal inference based on observational data. He argues that in epidemiology and the social sciences, qualitative analysis is indeed a “type of *scientific* inquiry” (italics added), within the framework of recognizing multiple types. In characterizing this form of qualitative analysis, Freedman emphasizes the contribution of “causal-process observations” (CPOs). In Freedman’s view, these strategically selected pieces of qualitative evidence play an essential role in disciplined causal inference. He comments pointedly that in the effort to advance knowledge,

> Progress depends on refuting conventional ideas if they are wrong, developing new ideas that are better, and testing the new ideas as well as the old ones. The examples show that qualitative methods can play a key role in all three tasks . . . . (Freedman 2010a: 232)

---

research design (though obviously, there is wide disagreement as to how these desired outcomes should be achieved, and whether in any given context they are achieved). Within this big-tent understanding of science, we are happy to be included in the tent.

Relatedly, Freedman underscores the fragility of much quantitative analysis.

> . . . far-reaching claims have been made for the superiority of a quantitative template that depends on modeling—by those who manage to ignore the far-reaching assumptions behind the models. However, the assumptions often turn out to be unsupported by the data…. If so, the rigor of advanced quantitative methods is a matter of appearance rather than substance. (Freedman 2010a: 232)

Against the backdrop of these starkly contrasting views of appropriate methods, in the second edition of *Rethinking Social Inquiry* (Brady and Collier 2010) we have added 125 pages of new text that focus on: (1) ongoing controversy regarding KKV's legacy; (2) continuing criticism of the standard quantitative template, including regression modeling and estimates of uncertainty; and (3) emerging arguments about both qualitative and quantitative methods that hold the promise of strengthening tools for causal inference.[4] The following pages provide an overview of these themes.

Ongoing Controversy over King, Keohane, and Verba

The methodological positions adopted by KKV continue to be of great importance in political science and well beyond. The book has an exceptionally high level of citations, and year after year it receives impressive sales rankings with online book sellers.

In the period since the publication of our first edition in 2004, quantitative and qualitative methodologists alike have underscored KKV's contribution. Philip A. Schrodt, a quantitative methodologist, argues that it has been the "canonical text of the orthodox camp" among political methodologists. In many graduate programs, it is considered "the complete and unquestionable

[4] A further initial point should be underscored. The focus in both editions of *Rethinking Social Inquiry* is on the study of causes and consequences—and specifically on causal inference. Of course, this focus represents just one facet of methodology. In our own work we have written extensively on conceptualization and measurement, and indeed, assessing causes and consequences emphatically calls for careful attention to concept formation and operationalization. Yet the central concern here is with causal inference.

truth from on high" (Schrodt 2006: 335). On the qualitative side, James Mahoney notes the book's striking importance and remarkable impact in political science (2010: 120).

Ironically, achieving "doctrinal status was not necessarily the intention of KKV's authors" (Schrodt 2006: 336), and their perspectives have doubtless evolved in the intervening years. Yet notably, in 2002—eight years after the book's original publication—King published an extended, programmatic statement on methodology, nearly the length of a short book, entitled "The Rules of Inference" (Epstein and King 2002). This publication departs little from the arguments of KKV.

KKV is controversial, as well as influential, and its continuing importance is of great concern to scholars disturbed by its narrow message. Our first edition already contained strong critiques, and new commentaries—some extremely skeptical—have continued to appear. These more recent arguments merit close examination.

Schrodt presents a bruising critique.

> KKV establishes as the sole legitimate form of social science a set of rather idiosyncratic and at times downright counterintuitive frequentist statistical methodologies that came together…to solve problems quite distant from those encountered by most political scientists…. (2006: 336)

Schrodt views KKV as promoting "a statistical monoculture" that is "not even logically consistent" (2006: 336). Indeed, he is convinced that

> one of the reasons our students have so much difficulty making sense of [KKV] is that in fact it does not make sense. (2006: 336)

Mahoney (2010), in his comprehensive essay "After KKV: The New Methodology of Qualitative Research," argues that KKV has "hindered progress in political science" by "controversially and perhaps unproductively promoting a singular quantitative approach" (2010:

121). Weyland, with obvious annoyance, suggests that the authors of KKV “offered to help out their inferentially challenged qualitative brethren,” proposing that their work should be “as similar as possible to quantitative studies.” The book in effect makes claims of “quantitative superiority” that “rest on problematic assumptions” (2005: 392), thereby reinforcing the mindset in which “qualitative research was often seen as lacking precision and rigor and therefore undeserving of the ‘methods’ label” (2005: 392).

In discussing the first edition of our own book, *Rethinking Social Inquiry*, Schrodt suggests that in this polarized context, “adherents of the [methodological] orthodoxy consider the heresies proposed therein to be a distraction at best; a slippery slope…at worst” (2006: 335). To take one example, what we would view as one of the orthodox commentaries is found in Nathaniel Beck (2006, 2010), who treats the idea of causal-process observations as an “oxymoron”—thereby essentially dismissing a basic concept in our book. He repeatedly acknowledges that scholars should “understand their cases” (e.g. 2006: 350) and that qualitative evidence contributes to this background knowledge, but he questions the idea that causal-process observations meet acceptable standards for causal inference (352).

Schrodt views elements of the response to *Rethinking Social Inquiry* among mainstream quantitative researchers as reflecting an unfortunate, defensive reaction. He argues that

> many in the statistical community have taken criticism of any elements of the orthodox approach as a criticism of all elements and circled the wagons rather than considering seriously the need for some reform. (Schrodt 2006: 338)

He also notes that when the editor of the methodology journal *Political Analysis* announced at the 2005 summer methodology meetings that the journal planned a symposium on *Rethinking Social Inquiry*, the room responded as if to express concern that “there are traitors in our midst!”

(2006: 338). Schrodt comments that this resistance reflects "a worrisome contentment with the status quo" among quantitative methodologists (2006: 338).

Based on this discussion, it seems clear that major controversies over methods stand behind these criticisms. We now explore two of these controversies.

Criticism of the Standard Quantitative Template

Statistical Modeling and Regression Analysis

In the past few years, the standard quantitative template centered on regression analysis has come under even heavier criticism. This development has two implications here. First, given KKV's reliance on this template, it further sharpens concern about the book's influence. Second, looking ahead, this development greatly extends the horizon of methodological approaches that should be—and in fact are being—discussed and applied, among both methodologists and users of alternative methods.

Much of this discussion centers on the enterprise of statistical modeling that stands behind regression analysis. In important respects, the precariousness of work with regression derives from the extreme complexity of statistical models. A statistical model may be understood as "a set of equations that relate observable data to underlying parameters" (Collier, Sekhon, and Stark 2010: xi). The values of these parameters are intended to reflect descriptive and causal patterns in the real world.

Constructing a statistical model requires assumptions, which often are not only untested, but untestable. These assumptions come into play "in choosing which parameters to include, the functional relationship between the data and the parameters, and how chance enters the model" (Collier, Sekhon, and Stark 2010: xi). Thus, debates on problems of regression analysis are simultaneously debates on the precariousness of statistical models. It is unfortunate that more than a few quantitative researchers believe that when the model is estimated with quantitative data and results emerge that appear interpretable, it validates the model. This is not the case.

We agree instead with the political scientist Christopher H. Achen, who argues that with more than two or three independent variables, statistical models will “wrap themselves around any dataset, typically by distorting what is going on” (2002: 443). Thus, what we might call a “kitchen sink” approach—one that incorporates numerous variables—can routinely appear to explain a large part of the variance without yielding meaningful causal inference. Relatedly, Schrodt states that with just small modifications in the statistical model, estimates of coefficients can

> bounce around like a box of gerbils on methamphetamines. This is great for generating large bodies of statistical literature…but not so great at ever coming to a conclusion. (2006: 337)

The econometrician James J. Heckman emphasizes that “causality is a property of a model,” not of the data, and “many models may explain the same data” (2000: 89). He observes that “the information in any body of data is usually too weak to eliminate competing causal explanations of the same phenomenon” (91).[5]

Sociologists have expressed related concerns, and Richard A. Berk concisely presents key arguments:

> Credible causal inferences cannot be made from a regression analysis alone…. A good overall fit does not demonstrate that a causal model is correct…. There are no regression diagnostics through which causal effects can be demonstrated. There are no specification tests through which causal effects can be demonstrated. (2004: 224)

[5] From the standpoint of econometrics, see also Leamer (1983, 36–38).

Berk amusingly summarizes his views in section headings within the final chapter of his book on regression analysis: “Three Cheers for Description,” “Two Cheers for Statistical Inference,” and “One Cheer for Causal Inference” (2004: chap. 11).[6]

Mathematical statisticians have likewise confronted these issues. Freedman’s skepticism about regression and statistical modeling has been noted above, and his incisive critiques of diverse quantitative methods have now been brought together in an integrated volume that ranges across a broad spectrum of methodological tools (Freedman 2010b).

Also from the side of mathematical statistics, Persi Diaconis argues that “large statistical models seem to have reached epidemic proportions” (1998: 797), and he laments the harm they are causing. He states that “there is such a wealth of modeling in the theoretical and applied arenas that I feel a sense of alarm” (804). Given these problems, methodologists should take more responsibility for the epidemic of statistical models by advocating “defensive statistics” (1998: 805). Thus, it should be a professional obligation to proactively warn scholars about the host of methodological problems summarized here.

In sum, many authors are now expressing grave concern about methods that have long been a mainstay of political and social science, and that are foundational for KKV.

Estimating Uncertainty

Standard practices in mainstream quantitative methods for estimating the uncertainty of research findings have also been challenged. The quest to estimate uncertainty is quite properly a high priority, prized as a key feature of good research methods. KKV views understanding and estimating uncertainty as one of four basic features of scientific research (1994: 9). In its

[6] Related arguments of sociologists have been advanced by Morgan and Winship (2007: passim) and Hedström (2008: 324). Statements by psychometricians include Loehlin (2004, 230–34).

discussion of "defining scientific research in the social sciences," the book states that "without a reasonable estimate of uncertainty, a description of the real world or an inference about a causal effect in the real world is uninterpretable" (9). The received wisdom on these issues is central to mainstream quantitative methods.

Unfortunately, KKV presumes too much about how readily uncertainty can be measured. In conjunction with the original debate over King, Keohane, and Verba, for example, Larry M. Bartels (2010) argues that these authors greatly overestimate the value of the standard insight that random error on an independent variable biases findings in knowable ways, whereas such error on the dependent variable does not. Bartels demonstrates that this would-be insight is incorrect.

A more pervasive problem involves significance tests. Any scholar acquainted with conventional practice in reporting regression results is well aware of the standard regression table with “tabular asterisks” scattered throughout.[7] The asterisks indicate levels of statistical significance, calculated on the basis of the standard errors of the coefficients in the table. Too often, when researchers report their causal inferences they simply identify the coefficients that reach a specified level of statistical significance. This is a dubious research practice.

A central problem here is that findings reported in regression tables are routinely culled from numerous alternative specifications of the regression model, which obviates the standard meaning and interpretation of the asterisks. Once again, Schrodt states the objection with particular clarity:

> The ubiquity of exploratory statistical research has rendered the traditional frequentist significance test all but meaningless. (2006: 337)

[7] Meehl 1978, cited in Freedman and Berk 2010: 24.

Freedman and Berk (2010: 24) underscore the dependence of significance tests on key assumptions. For descriptive inference (external validity), they assume a random sample, rather than the convenience sample common in political science. Even with a random sample, missing data—including the problem of non-respondents—can make it more like a convenience sample.[8] Another assumption requires a well-defined—rather than ill-defined or somewhat arbitrarily defined—population.

For causal inference (internal validity), avoiding data snooping is crucial if significance tests are to be meaningful. Here, the presumption is that the researcher has begun with a particular hypothesis and tested it only once against the data, rather than several times, adjusting the hypothesis and model specification in the search for results deemed interesting. This inductive approach is *definitely* a valuable component of creative research, but it muddies the meaning of significance tests.

Against this backdrop, Freedman, Pisani, and Purves (2007) are blunt in their warnings:

1. "If a test of significance is based on a sample of convenience, watch out" (556).
2. "If a test of significance is based on data for the whole population, watch out" (556).
3. "Data-snooping makes P-values hard to interpret" (547).
4. "An 'important' difference may not be statistically significant if the N is small, and a unimportant difference can be significant if the N is large" (553).[9]

A key point should be added. In his various single-authored and co-authored critiques of significance tests, Freedman specifically does not turn to the alternative of Bayesian analysis. Rather, as in his other writings on methodology, he advocates common sense, awareness that

[8] See Freedman (2008b: 15). Thus, starting with a random sample, in the face of problems such as resource constraints that limit tracking down respondents, the researcher can end up with what is in effect a type of convenience sample.
[9] I.e., if assumptions are not met, "significance" level depends on the sample size, without reflecting the real meaning of statistical significance.

statistical tools have major limitations, and substantive knowledge of cases as an essential foundation for causal inference. [10]

Where Do We Go from Here?

The practical importance of these problems is quickly seen in the fact that, to a worrisome degree, a great deal of quantitative research in political science has proceeded as if regression-based analysis, including accompanying measures of uncertainty, yields reliable causal inference. A vast number of journal articles have sought to make causal inferences by estimating perhaps half a dozen related (and commonly under-theorized) model specifications, picking and choosing among these specifications, and offering what is too often an ad hoc interpretation of a few selected coefficients—generally, quite inappropriately, on the basis of significance levels. These failings have been further exacerbated by the readily available statistical software that makes it easy for researchers with virtually no grasp of statistical theory to carry out complex quantitative analysis (Steiger 2001).

In the face of these serious challenges, we advocate two avenues of escape: refinement of qualitative tools, and further innovation in quantitative methods. These alternatives are explored in the new chapters of *Rethinking Social Inquiry*, second edition.[11]

Qualitative Tools

An important avenue is opened by further refinements in qualitative analysis. One familiar, traditional option here is the small-N comparative method, a strategy common both in cross-national comparisons, and also in comparisons of political units within nations—potentially involving regions, provinces or states, or metropolitan areas. Here, the analyst

[10] See, for example, Freedman 2008, 2010a.
[11] Four chapters are entirely new. Henry Brady's analysis of data-set observations versus causal-process observations was an appendix in the first edition.

juxtaposes perhaps two, or four, or six units, selecting matching and contrasting cases so as to "control" for extraneous factors and allow a focus on the principal variables of concern. This approach is often identified with J. S. Mill's (1974 [1843]) methods of agreement and difference, and Przeworski and Teune's (1970) most similar and most different systems designs.

In our view, this small-N comparative approach is invaluable for concept formation, pinning down indicators, and discovering and formulating explanatory ideas. However, it is much weaker as a basis for causal inference, given that it involves what is in effect a correlation analysis with a very small N. The matching and contrasting of cases employed cannot by itself succeed in controlling for variables the researcher considers extraneous to the analysis. Any presumption that this matching of cases creates a natural experiment or a quasi-experiment is misleading.

Rather, the key step is to juxtapose this comparative framing with carefully-executed research carried out within the cases, quite often involving qualitative analysis. At this point, the tools of causal-process observations and process tracing become crucial. They certainly do not solve all problems of qualitative analysis, yet they make valuable contributions.

Regarding the first of these, a central element in our effort to place qualitative analysis on a more rigorous foundation (Collier, Brady, and Seawright 2010) has been our distinction between: (1) data-set observations (DSOs), which correspond to the familiar rectangular data set of quantitative researchers; and (2) causal-process observations (CPOs), i.e., pieces of data that provide information about context, process, or mechanism and contribute distinctive leverage for causal inference in qualitative research.[12]

[12] In quantitative research, the idea of an "observation" (as in DSO) has special status as a foundation for causal inference, and we deliberately incorporated this label in the idea of CPOs to underscore their relationship to causal inference. We are pleased that other scholars have also found the idea of CPOs to be useful. For example, see

Causal-process observations and the widely-discussed procedure of "process tracing" are closely connected. When *process tracing* is used for causal inference,[13] the pieces of evidence on which the researcher focuses are specifically *causal-process observations*. Process tracing consists of procedures for singling out particular CPOs as are relevant for causal inference in a given context.

In our second edition, Andrew Bennett, David A. Freedman, and Henry E. Brady explore different facets of causal-process observations and process tracing.

Bennett's (2010) chapter provides a new introduction to process tracing, which is routinely invoked as a basis for causal inference in qualitative work. Unfortunately, the specific steps involved are often poorly understood—possibly one reason why many quantitative researchers are skeptical about causal inference in qualitative studies.

Bennett formulates a typology that places on two dimensions the alternative tests employed in process tracing. The tests are distinguished according to whether passing a particular test is *necessary* for inferring causation, and whether it is *sufficient*. This provides a new framework for thinking about the tests originally proposed by Van Evera (1997: 31–32): the straw-in-the wind, hoop, smoking-gun, and doubly-decisive tests. Bennett illustrates his framework by applying the typology at the level of macro-politics, focusing on three well-known historical episodes in international relations.

Next, the statistician David Freedman's (2010a) chapter examines the role of causal-process observations in six major studies from the history of epidemiology, including John Snow's famous research on the causes of cholera. In Freedman's view, both epidemiology and the social sciences use qualitative evidence as a basic type of scientific inquiry, which he

Mahoney 2010: 123–131, who uses it as one of the organizing principles in his synthesis of new trends in qualitative methods. The concept is also employed by Freedman (2010a).

[13] Mahoney 2010: 125–28 explains how it is also used for descriptive inference.

believes is often more productive than conventional quantitative approaches. More than a few quantitative researchers need to examine Freedman's views and reconsider their skepticism about inference in qualitative research.

Freedman is very specific about the contributions of qualitative evidence, arguing that it can play a valuable role in overturning prior hypotheses, as well as formulating and testing new hypotheses. In this sense, Freedman's position is quite different from that of Fearon and Laitin (2008: 756), who sharply subordinate qualitative vis-à-vis quantitative analysis. It also challenges Piore's (2006: 17) assertion that information from case studies "cannot be treated directly as empirical evidence…."

Yet Freedman is also strongly committed to the careful juxtaposition of CPOs and DSOs. He is skeptical about much quantitative analysis, and he prefers quantitative research—for example, natural experiments—that is carried out jointly with careful qualitative analysis.

Finally, Henry Brady's (2010) chapter, focused on electoral behavior, shows how a sequence of CPOs can yield causal inference. This procedure gives crucial leverage in disputing the claim, advanced in a quantitative study by John Lott, that George Bush lost thousands of votes in the 2000 presidential election in Florida due to the media's early (and incorrect) call of the election outcome. Brady's working hypothesis is that the early call had little impact on the actual vote. He goes through a series of process-tracing steps to support his hypothesis, employing a sequence of vote counts and assumptions about voting behavior that identified the necessary conditions for Lott's hypothesis to be plausible.

Brady's analysis reminds us of the obvious but crucial point that process tracing can employ numerical data. The use of such data does not necessarily involve DSOs, i.e., a rectangular data set. Rather, isolated pieces of quantitative information are treated as CPOs.

Looking beyond these three chapters, we believe—notwithstanding extensive efforts to institutionalize graduate training in qualitative methods[14]—that this training does not adequately address process tracing and causal-process observations. This deficit has motivated us to prepare a set of exercises, available online, for teaching these analytic tools.[15]

Quantitative Methods

The final two chapters in the second edition consider further the problems with quantitative methods and potential solutions to these problems. Jason Seawright's (2010) "Regression-Based Inference: A Case Study in Failed Causal Assessment" examines a number of unsuccessful attempts to employ quantitative, cross-national regression analysis to address a classic topic in comparative social science: the impact of political regime type on economic growth.

Seawright's substantive focus is of great importance, and his analysis has broad implications for the vast literature that uses the quantitative, cross-national method to study dozens of major topics in political and social science. Indeed, in discussing these failures of inference, he cites authors who call for an "obituary" for a significant part of quantitative cross-national literature. His arguments are also relevant for more general discussions of regression analysis. For example, many scholars believe they can improve regression-based inference by simply adding more control variables. Yet introducing further controls can potentially make inferences worse, rather than better. Seawright goes on to show that refinements in regression—such as designs employing matching and instrumental variables—are at best problematic in addressing these problems. In conclusion, he proposes that scaling down to a finer-grained focus,

[14] Notable among these efforts, the APSA Organized Section for Qualitative and Multi-Method Research (which publishes this newsletter) sponsors short courses at the annual political science meetings, as well as numerous panels of great pedagogical value. The annual Institute for Qualitative and Multi-Method Research at Syracuse University also provides excellent instruction. These initiatives are discussed in Collier and Elman (2008).
[15] "Exercises for Teaching Process-Tracing" are found at the top of David Collier's UC Berkeley website <polisci.berkeley.edu/people/faculty>.

which may include substantial use of qualitative analysis, can come to the rescue in overcoming these failures.

The final chapter by Thad Dunning (2010) evaluates efforts at "design-based inference." He focuses on the family of techniques known as natural experiments, including regression-discontinuity and instrumental variable designs. These techniques seek to overcome the failures of regression explored by Seawright and discussed earlier in this article. Dunning creates a typology that assesses these designs on three dimensions: (1) plausibility of the "*as-if* random assignment," an assumption central to natural experiments; (2) credibility of the statistical model; and (3) substantive relevance of the key explanatory variable. This third point is especially important because some of these studies, in the search for situations of *as-if* random assignment, drastically limit the substantive relevance of their investigations. Indeed, this shortcoming is a serious weakness of natural experiments.

By juxtaposing the three dimensions in his typology, Dunning evaluates whether particular natural experiments employ strong or weak research designs. He identifies a small number of unusually successful natural experiments, yet he also shows that many famous studies in this tradition do poorly in terms of his criteria. Hence, although Dunning views natural experiments as a promising tool, he argues that they should definitely not eclipse other methodologies.

Finally, Dunning underscores repeatedly the importance of qualitative data and qualitative insight in the design and execution of natural experiments. Without this foundation, researchers lack a basis for many judgments and decisions essential to the method—for example, the key assumption of as-if random assignment. Hence, this newly popularized methodology is strongly dependent on traditional tools—qualitative evidence, hard-won insight into the details of cases, and knowledge of context. It is a striking example of why a multi-method approach is invaluable in good research.

* * * * * * * *

To summarize: We are unquestionably observing a sea change in political methodology. Conventional quantitative methods are now the focus of even sharper criticism. The tools of qualitative analysis are being further refined and legitimated in ways that addresses some of these failings. These qualitative tools certainly have their own problems and limitations, but we think real progress is being made. Natural experiments likewise show promise, though they are far from being a methodological Nirvana. Correspondingly, in our view an eclectic practice of methodology—involving the idea of "diverse tools" highlighted in the title of our book—is the most promising avenue to pursue.

References

Achen, Christopher H. 2002. “Toward a New Political Methodology: Microfoundations and ART.” *Annual Review of Political Science* 5: 423–450.

Bartels, Larry M. 2010. “Some Unfulfilled Promises of Quantitative Imperialism.” Chap. 6 in Brady and Collier, *Rethinking Social Inquiry*, 2nd edn.

Beck, Nathaniel. 2006. “Is Causal-Process Observation an Oxymoron?” *Political Analysis* 14:3 (Summer), 347–352.

Beck, Nathaniel. 2010. “Casual-Process ‘Observation’: Oxymoron or (Fine) Old Wine.” *Political Analysis* 18:4 (Autumn), 499–505.

Bennett, Andrew. 2010. “Process-Tracing and Causal Inference.” Chap. 10 in Brady and Collier, eds., *Rethinking Social Inquiry*, 2nd edn.

Brady, Henry E. 2010. “Data-Set Observations versus Causal-Process Observations: the 2000 U.S. Presidential Election.” Chap. 12 in Brady and Collier, *Rethinking Social Inquiry*, 2nd edn.

Brady, Henry E., and David Collier, eds. 2004/2010. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 1st and 2nd edns. Lanham, MD: Rowman and Littlefield.

Collier, David, and Colin Elman. 2008. “Qualitative and Multi-Method Research: Organizations, Publication, and Reflections on Integration.” In Janet Box-Steffensmeier, Henry E. Brady, and David Collier, eds., *Oxford Handbook of Political Methodology*. Oxford: Oxford University Press.

Collier, David, Henry E. Brady, and Jason Seawright. 2010. “Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology.” Chap. 9 in Brady and Collier, eds., *Rethinking Social Inquiry*, 2nd edn.

Collier, David, Jasjeet Sekhon, and Philip Stark. 2010. “Editors’ Introduction: Inference and Shoe Leather.” In David A. Freedman, *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*, David Collier, Jasjeet Sekhon, and Philip B. Stark, eds. Cambridge: Cambridge University Press.

Diaconis, Persi. 1998. “A Place for Philosophy? The Rise of Modeling in Statistical Science.” *Quarterly of Applied Mathematics* 56:4 (December), 797–805.

Dunning, Thad. 2010. “Design-Based Inference: Beyond the Pitfalls of Regression Analysis?” In Brady and Collier, eds., *Rethinking Social Inquiry*, 2nd edn.

Epstein, Lee, and Gary King. 2002. “The Rules of Inference.” *University of Chicago Law Review* 69:1 (Winter), 1–133.

Fearon, James D., and David D. Laitin. 2008. “Integrating Qualitative and Quantitative Methods.” In Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, eds. *Oxford Handbook of Political Methodology*. Oxford: Oxford University Press.

Freedman, David A. 2008. “Rejoinder.” *Newsletter of the American Political Science Association Organized Section for Qualitative and Multi-Method Research* 6:2 (Fall), 14–16.

Freedman, David A. 2010a. “On Types of Scientific Inquiry: The Role of Scientific Reasoning.” In Henry E. Brady and David Collier, eds., *Rethinking Social Inquir*y, 2nd edn.

Freedman, David A. 2010b. *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*, David Collier, Jasjeet Sekhon, and Philip B. Stark, eds. Cambridge: Cambridge University Press.

Freedman, David A., Robert Pisani, and Roger Purves. 2007. *Statistics*, 4th edn. New York: Norton.

Freedman, David A., and Richard A. Berk. 2010. “Statistical Assumptions as Empirical Commitments.” In David A. Freedman, *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*, David Collier, Jasjeet S. Sekhon, and Philip B. Stark, eds. New York: Cambridge University Press.

Heckman, James J. 2000. “Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective.” *Quarterly Journal of Economics* 115: 1 (February), 45–97.

Hedström, Peter. 2008. “Studying Mechanisms to Strengthen Causal Inferences in Quantitative Research.” In Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, eds. *Oxford Handbook of Political Methodology*. Oxford: Oxford University Press.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Leamer, Edward E. 1983. “Let’s Take the Con out of Econometrics.” *American Economic Review* 73:1 (March), 31–43.

Loehlin, John C. 2004. *Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis.* Mahwah: Lawrence Erlbaum Associates, Publishers.

Mahoney, James. 2010. “After KKV: The New Methodology of Qualitative Research.” *World Politics* 62:1 (January), 120–147.

Meehl, Paul E. 1978. “Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald, and the Slow Progress of Soft Psychology.” *Journal of Consulting and Clinical Psychology* 46:4 (August), 806–834.

Mill, John Stuart. 1974 [1843]. “Of the Four Methods of Experimental Inquiry.” *A System of Logic*. Toronto: University of Toronto Press. Originally published in 1843.

Morgan, Pamela S. 1996. “‘What Is Science?’ and the Cognitive Theory of Contested Concepts.” Department of Linguistics, University of California, Berkeley.

Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research.* New York: Cambridge University Press.

Piore, Michael J. 2006. “Qualitative Research: Does it Fit in Economics?” *European Management Review* 3:1 (Spring), 17–23.

Przeworski, Adam, and Henry Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley.

Schrodt, Philip A. 2006. “Beyond the Linear Frequentist Orthodoxy.” *Political Analysis* 14:3 (Summer), 335–339.

Steiger, James H. 2001. “Driving Fast in Reverse.” *Journal of the American Statistical Association* 96:453 (March), 331–338.

Weyland, Kurt. 2005. “Book Review.” *Perspectives on Politics* 3, (June), 392–93.