

Qualitative Methods

Newsletter of the
American Political Science Association
Organized Section on Qualitative Methods

Contents

Symposium: Ethnography Meets Rational Choice: David Laitin, For Example

| | |
|---|----|
| <i>Introduction</i> | |
| Ted Hopf | 2 |
| <i>Theory, Data, and Formalization: The Unusual Case of David Laitin</i> | |
| Yoshiko M. Herrera | 2 |
| <i>Mechanisms vs. Outcomes</i> | |
| Kanchan Chandra | 6 |
| <i>Understanding Rules and Institutions: Possibilities and Limits of Game Theory</i> | |
| David M. Woodruff | 13 |
| <i>Ethnography and Rational Choice in David Laitin: From Equality to Subordination to Absence</i> | |
| Ted Hopf | 17 |
| <i>Recognizing the Tradeoffs We Make</i> | |
| Ashutosh Varshney | 20 |
| <i>Ethnography and/or Rational Choice: A Response from David Laitin</i> | |
| David Laitin | 26 |

Symposium: Alexander L. George and Andrew Bennett's Case Studies and Theory Development in the Social Sciences

| | |
|--|----|
| <i>Introduction</i> | |
| Jack S. Levy | 34 |
| <i>Notes From a Generalist</i> | |
| Daniel W. Drezner | 34 |
| <i>A Major Milestone with One Major Limitation</i> | |
| John S. Odell | 37 |
| <i>Well Worth the Wait</i> | |
| Jack S. Levy | 40 |
| <i>Case Studies and the Philosophy of Science</i> | |
| David Dessler | 43 |
| <i>Advancing the Dialogue on Qualitative Methods</i> | |
| Andrew Bennett | 45 |
| Book Notes | 48 |
| Announcements | 51 |

Letter from the Editor

John Gerring
Boston University
jgerring@bu.edu

In the Fall of 2002, APSA created its 37th Organized Section, devoted to the study, development, and dissemination of qualitative methods. Since that time, I have served as the editor of this newsletter. My job, as I saw it, was to bring to the attention of our members the most interesting, innovative, and (it follows) contentious issues in the field of political methodology, regardless of whether they might be categorized conventionally as 'qualitative' or 'quantitative.' (Issues of import solely to quantitative work have been deferred to the Political Methodology section—no need to duplicate effort.) With that caveat, the mission of the newsletter was interpreted broadly to include all methodological issues of relevance to the study of politics. Symposia have ranged from broad philosophy-of-science issues to narrower debates about technique. For the most part, these topics have been chosen in response to ideas from our members and as extensions of APSA panels and roundtables. Usually, the management of a symposium was delegated to the person taking the initiative to organize a discussion on that topic.

As editor I took a *laissez-faire* approach to the newsletter, asking authors to follow only a few stylistic and substantive guidelines: contributions should be short, accessible to a broad readership, written with some flair, and encompassing a range of viewpoints on the chosen subject. The aim, while retaining some of the intellectual rigor associated with more traditional academic journals, was to give writers scope to opine—that is, to use the first-person pronoun and to adopt a more discursive manner than would be usual in a more formal academic venue. In this manner, I hoped to reproduce the lively and candid views exchanged with each other in emails and over cups of joe. "What do you really (in your heart of hearts) think of X?" This is the sort of conversation that I wanted to foster.

During the last three years, the newsletter has covered a lot of ground. In Spring 2003, we ran a symposium on teaching qualitative methods, which featured a comprehensive review of textbooks and discussions of courses and various approaches to the subject. In Fall 2003, our symposium addressed the knotty issue of "interpretivism," with contributions from several scholars, including Clifford Geertz. In Spring

2004, we ran two symposia, the first on techniques of field research and the second on content and discourse analysis. In Fall 2004, we tackled Charles Ragin's complex and innovative technique of Qualitative Comparative Analysis (QCA), and its recent variants. In Spring of last year we ran symposia on the qualitative/quantitative distinction (with contributions from both sides of the divide), and on the use of necessary-condition causal propositions. This past Fall we featured a discussion of where new hypotheses originate, by Richard Snyder, along with two symposia, one focused on Ian Shapiro's *The Flight from Reality in the Human Sciences*, and the second devoted to the subject of concept formation in the social sciences.

Every year we take notice of recent methodological publications that may be of interest to our readers. (For the reasons mentioned above, we don't cover work that is narrowly tailored to statistical analysis.) The *Book Notes* and *Article Notes* features are intended to list work that either has an explicit methodological focus or uses an innovative technique to good effect. If you know of a book or article published since 2000 that has not already appeared in these pages—and has a strong methodological theme or innovation—do let us know. (Self-nominations are encouraged!)

In this issue, we are fortunate to be able to feature two roundtables focused on the work of scholars who have had

enormous influence on the discipline. The first examines the career of David Laitin, whose work incorporates ethnography and rational choice—methods often deemed to be antithetical—and the second solicits comments on the recent landmark publication by Alexander L. George and Andrew Bennett, *Case Studies and Theory Development* (MIT Press, 2005).

The following issue (Fall 2006) will begin the tenure of a new editor, whom I am delighted to introduce. Gary Goertz has written widely on international relations and on methodological issues and teaches regularly at IQRM, the winter graduate training institute at Arizona State University. Having engaged with both quantitative and qualitative methodological issues, he is well positioned to foster a productive debate among scholars who utilize diverse approaches to the study of politics. I know that Gary is looking forward to engaging with the QualMeth community and wishes to hear your ideas on how to maintain the newsletter as a vital part of our research community. Please join me in welcoming Gary, and please accept my thanks for your participation in the newsletter's ongoing activities. Finally, let me take this opportunity to thank Joshua Yesnowitz, who has served as our assistant editor for the past several years and will continue under Gary's tenure. Josh has done a superb job of keeping track of the details and putting everything together. We are grateful for his stewardship.

Symposium: Ethnography Meets Rational Choice: David Laitin, For Example

Introduction

Ted Hopf

Ohio State University
hopf.2@osu.edu

One of the more encouraging developments in political science over the last few years has been the appearance of work that is self-consciously multi-methodological. An increasing number of dissertations and publications combine formal models with statistical analysis of large-n data sets and comparative case studies.

Less evident are efforts to combine ethnography, or the recovery of the intersubjective world of actors themselves, with more mainstream traditional or formal methods. David Laitin is one of the rare scholars who has engaged in serious ethnography (*Hegemony and Culture*), combined ethnography with other methods (*Identity in Formation*), and applied rational choice techniques, with James Fearon, to issues of identity. ("Explaining Interethnic Cooperation") His work provides the opportunity for this symposium.

Each of the authors has critically engaged Laitin's work, with an eye toward assessing the merits and possibilities of combining serious ethnographic scholarship with rational choice. While conclusions are best left to readers themselves, it is fair to say that the authors share concerns with how eth-

nographic sensitivity to contextual realities can be squared with the a priori simplifications necessitated by rational choice approaches. But, importantly, each of the authors also believes it is a combination well worth attempting.

Each of the papers in this symposium was originally presented as a Qualitative Methods Roundtable at the September 2005 American Political Science Association meetings in Washington. David Laitin's responses to these papers concludes this symposium, but begins a long, continuing conversation with his many critical admirers.

Theory, Data, and Formulation: The Unusual Case of David Laitin

Yoshiko M. Herrera

Harvard University
herrera@fas.harvard.edu

In two influential articles David Laitin laid out a tripartite method for comparative politics and for social science more generally (Laitin 2002, 2003). The three methods that Laitin advocated were Formal Theory, Quantitative Analysis, and Narrative. In this paper I take issue with Laitin's categorization scheme for the methods, and I consider the criteria and constraints on choosing methods.

Types of Methods

Laitin's tripartite framework is an intriguing methodological model for the social sciences. While I agree with many aspects of the argument, in contrast to Laitin's tripartite framework I believe there are actually two major distinctions separating the vast array of methodologies available to social scientists: theory vs. data, and formalization vs. non-formalization. Moreover I think that these distinctions might usefully form the basis of a two-by-two categorization scheme yielding four types of methods: Theory, Formal Theory, Narrative, and Quantitative Analysis (see Table 1).

Table 1

| | Analysis of Ideas | Analysis of Data |
|-------------------|--------------------------|----------------------------|
| Non-formal | Theory (T) | Narrative (N) |
| Formal | Formal Theory (FT) | Quantitative Analysis (QA) |

The primary difference in my formulation from Laitin's framework is the distinction between methodologies that focus on the analysis of data and those that focus on the analysis of ideas or theories, and the division of theory into formal and non-formal types.

The four methodological *labels*—Theory, Formal Theory, Narrative and Quantitative Analysis—stand in of course for a wide array of actual methods. Non-formal theory encompasses any abstract thought, philosophy, or set of rules, principles, beliefs, or ideas which has not been formalized into mathematical language. That which has been formalized into mathematical concepts, including social choice, game-theory, differential equation modeling, etc., can be called formal theory. Formal theory and theory are fundamentally about the analysis of ideas rather than data.

In contrast, narrative and quantitative analysis share an analytic focus on data, rather than ideas. These data can be derived from a variety of sources (interviews, texts, surveys, etc.), but what separates narrative from quantitative analysis is the formalization of data into quantified or numerical entities subject to statistical methods. Narrative can mean ethnography, discourse analysis, case studies, or any analysis of data that has not been formalized through quantification. Similarly, quantitative analysis is the examination of quantified data using a variety of statistical methods from simple significance tests in cross-tabulations to a wide range of regression models including OLS, probit, and Bayesian statistics.

Laitin did emphasize the distinction between formal theory and quantitative analysis, but I think his tripartite framework did not go far enough in differentiating the two. We must push understanding of the differences further. In practice there appears to be an affinity between quantitative analysis and formal theory, but it's crucial to properly understand what the two methods share, as well as what distinguishes them. Formal theory and quantitative analysis primarily share one thing:

mathematical language. That is, they are both formal in the sense of relying on mathematics to work out complicated relationships among variables, and they accept the constraints imposed by that language. This shared language accounts for why it might sometimes seem that quantitative analysis scholars and formal modelers are better able to talk to each other than those who don't share their language, i.e. those working in non-formal theory or methods. But, a shared language, or shared formalization, does not bridge the ocean of difference between the two methods in terms of the object of analysis, namely theory vs. data. Quantitative analysis is fundamentally about data. It uses the science of statistics to manipulate empirical evidence. Often the goal is to empirically test existing theories, either formal or non-formal. But without data, there is no quantitative analysis.

Quantitative types hunger for large data sets, and without data sets, the operation comes to a halt. Indeed, there are those so committed to QA that if the datasets do not exist for a given problem, they do not study it; and there are some who will resort to using any data, no matter how poor, as long as they exist. The key point is that QA shares with narrative—or history, or ethnography, or discourse analysis—an analytic focus and dependence on empirical data.

These data-focused methodologies are very different from formal modeling or theory. Formal theory is fundamentally a theoretical exercise. The common substitution of *theory* for *modeling* in the title is not a coincidence. Formal theorists do not set out first and foremost to solve empirical puzzles; rather, they make their living formalizing ideas that have not yet been formalized. In confronting a topic, the formal theorist asks, has this been modeled? Or, in other words, has the logic or set of ideas been put into mathematical language? To put it more starkly, formal theory can operate without datasets, and in a world without datasets, the formal theory enterprise would hardly suffer. Ideas, not data, are the foundation of theory, both formal and non-formal. Thus, formal theory shares much with non-formal theory, and it is the analysis of ideas that links these two methods and separates them from both quantitative analysis and narrative.

How Many and Which Methods to Use?

What has turned out to be most provocative about Laitin's tripartite method was not the list of methods themselves or the categorization scheme which included the three methods, but the idea which many people have taken from the discussion of the tripartite method, namely that all social scientists should use all three methods together in all of their research projects.

At the 2005 APSA panel, Laitin argued that this was a misinterpretation of the tripartite method, and that his view was that the three methods should be used collectively by social scientists in a way such that for any given research problem all three methods get employed, and practitioners of each method appreciate in their own work the contributions of the other methods. Indeed, in the 2002 chapter, Laitin writes, "my argument is not that all comparativists should have highly cultivated statistical, formal, and narrative skills" (Laitin 2002,

659) and he goes on to argue that no one method should dominate the discipline. However, if one reads the 2002 article carefully, we see that Laitin does not actually say how many methods one person should use; he suggests three is too many for one person and he argues for collective diversity, but he then leaves it at that.

Rather than clarifying, in the 2003 article, Laitin starts out remarkably vague on the issue of just how many methods one researcher *ought* to use.¹ In critiquing two scholars, Bent Flyvbjerg and Stanley Tambiah, whose primary contributions are to the use of narrative methods, he writes, “the work would have much greater scientific value if placed within what I have dubbed the *tripartite method of comparative research*—a method that integrates narrative..., statistics, and formal modeling” (Laitin 2003, 164-5, emphasis in original). But what does it mean to say the narrative work should be “placed within” the tripartite method? Does it mean that Flyvbjerg and Tambiah should have added quantitative and formal modeling to their analyses? Or does it mean some other scholars should have come along and studied the same problems as Flyvbjerg and Tambiah using other methods? There is an enormous difference between saying social science research problems should be studied from a number of methodological angles by different people and saying individual researchers should use all of the methods.

Farther on in the article, Laitin suggests that he means the latter, namely that individuals should use all three methods. Laitin calls Randall Stone’s work, which uses all three methods, “an exemplary model of the tripartite method” (Laitin 2003, 177). It is interesting that Laitin did not choose a collective enterprise of several scholars working on the same problem using different methods as the “exemplary model” of the tripartite method, but rather he chose an individual who had used all three methods. This choice suggests that the view that Laitin is advocating the use of all three methods by individuals is not outside the range of reasonable interpretation.

Whether or not Laitin was actually arguing for the use of multi-methods by individuals or merely advocating collective diversity (or some of each), his writings in any case still highlight the question of how many methods individual researchers should be expected to use in one project. I argue that the four methods in the framework proposed above, or even the three methods proposed by Laitin, are not necessary or desirable in every social science project. In my view, the appropriateness or choice of methods depends on four factors:

- (1) the nature of the problem under investigation, and the contributions that a particular method might make to such a study;
- (2) the resources available to a scholar;
- (3) the disciplinary context, including norms, incentives and constraints, in which a scholar works; and
- (4) the aptitude and will of a scholar to do multi-method work.

No one method is a panacea for all social science problems. In particular, it is a mistake to assume that formal theory is the only type of theory useful to social science. Many for-

mal theorists would agree that formalization of a theoretical argument is not always required or advantageous. Sometimes a problem is so simple as to make formalization redundant. Sometimes it’s been done before, and therefore another formalization would yield no new insights. And most importantly, there are classes of problems for which formalization is not appropriate (yet, or perhaps ever): these include issues such as irrational beliefs and behavior, non-transitive preferences, and interactions in which new unknown and unknowable possibilities for action exist. Formal theory is not able to solve these sorts of problems at the moment, and may never be able to. While formal theorists are currently working on expanding the range of formal models, it is very unlikely that in the near future all theoretical issues will be subject to formal analysis. Thus, sometimes a case can be made for formalization of theory, and sometimes not.

In addition to the specificity of the problem at hand, resources are another factor for scholars in choosing the appropriate set of methods. Methods are costly to learn and to do in practice, and therefore the more methods one chooses, the more costly it is for each researcher. Sometimes there are public goods such as data sets or publicly available empirical material, or existing theories or formal models which can be built upon, that lower the cost for researchers using a particular method, but this is not always the case, and thus, often, choosing methods requires trade-offs, which in practical terms might mean leaving out a method, despite its potential benefit. Hence the following types of contributions by individual scholars: theoretical work or formal models which have no empirical component, or work which uses only narrative or quantitative analysis rather than both. That these are the contributions of individuals does not mean that others cannot add to the collective contribution by bringing additional methods to the study of the same problem, but resource constraints do and probably will continue to limit the number of methods that any individual can employ.

Discipline or sub-disciplinary norms also play a role in which methods researchers choose, and there is variation across fields. For example, if we survey the social sciences as a whole, it is economics and political science, rather than anthropology, for example, where formal theory has been most successful. And in political science, it is the subfields of IR and American politics where formal modeling has advanced furthest. The reasons for this differential success are related to issues of problem-appropriateness and resources discussed above, but also, I believe, to the expectations and receptivity of different subfields to theoretical versus empirical work. Fields where theory is privileged and case-specific data requirements are the least rigorous are most receptive to theoretical enterprises such as formal modeling. In political science, comparative politics arguably has higher case-specific data requirements than IR. Comparativists, like anthropologists for example, are expected to know a lot about a given place, and to have experience in collecting data in that place. The comparativist who does not speak the language of his or her region of expertise is unlikely to get a job. Not so in IR. The IR area-specialist who works with regional documents in

original languages is a rare species. Most IR scholars are generalists, working on a topic in many places, without deep, language-dependent knowledge of particular places. This is no criticism, it's just an observation, which I think suggests that IR as a field is less demanding of detailed case-specific data than comparative politics. This may partially explain why formal theory has made more contributions in IR than comparative politics. It is not to say that formal theory is incompatible with case-specific data or fields that rely on primary language or case-specific data, but just that as a costly method which occasionally leaves out empirical work, it is not so surprising to find it has had a more positive reception in IR than in comparative politics.

Resource constraints or availability of data and theory also interact with disciplinary norms. In fields where empirical data or datasets are readily available, and therefore the cost of providing empirical tests are lower, there tends to be a higher threshold for expectations of empirical work. For example, in American politics, the sheer number of people working in the subfield and the accumulated body of both theory and empirical evidence in narrative as well as quantitative form mean that it is much less costly for a scholar to include multiple methods in a project than it would be for someone studying a place with much less existing work, theoretical and empirical, such as Sri Lanka. Thus, is it not so surprising that on average Americanists probably employ, individually, the greatest number of methods on any given project.

Finally, disciplinary norms also effect the type of the problems most likely to be studied in particular subfields, and therefore also have an effect on the appropriateness of particular methods. Fields that focus on individual actors or unitary actors are most receptive to methodological individualism and game-theoretic models of strategic interaction common in formal modeling. While these types of problems are common in many subfields of political science, they may be most common in IR, where, not-coincidentally, formal modeling has become so widely established.

To use all four methods—theory, formal theory, narrative, and quantitative analysis—is a nice proposition to consider as an ideal type, but as the examples above demonstrate, in practice it is rarely feasible owing to the nature of the problems under investigation, resources, and disciplinary contexts. Similarly, even in the case of Laitin's tripartite method, the use of all three methods is very uncommon in practice. A telling illustration of this point is that in his discussion in support of his tripartite method, Laitin could not find three good examples in the political science literature (though perhaps he restricted the pool to IR and comparative politics). Recall that he cited Robert Bates' coffee study, which used formal theory and narrative (Bates 1998); then Adam Przeworski et al.'s work on democracy, which used quantitative analysis and narrative (Przeworski et al. 2000); and finally Randall Stone's work, which did use all three methods (Stone 2002), but which Laitin criticized for the inadequate contribution of the narrative component, claiming it was overshadowed by the quantitative analysis and the formal model. Ironically, had Laitin cited his own work, he would have had examples of the tripartite method,

but to do so would have highlighted how rare such choices are outside of American politics. This point brings up the final factor in choosing methods, which is the will of individual researchers.

David Laitin's methodological breadth is not common. Over the course of his career he has not just learned new methods, but also new languages and new places to test his theories. This constant retooling is both costly and rare. After a certain point—often middle age—most scholars are not interested in moving out of their established comfort zone. Area specialists stick with their area; and quantitative people stick with quantitative analysis, for example. Sometimes people innovate on the margins: scholars working on narrative case studies expand their data analysis techniques and move into quantitative analysis, or quantitative scholars add case studies or archival material, or formal modelers expand their mathematical knowledge to move into quantitative analysis.

But David Laitin has gone far beyond these types of marginal innovations. He began his lifelong work on ethnic and language politics using narrative methods on African cases. He followed ethnic politics to Europe and Catalonia, learning new languages and places. But the end of the USSR seemed to provide a major experimental testing ground for language and identity politics so he again learned new languages, Russian and Estonian, and also new quantitative methods (content analysis and experiments), and began his study of formal theory. His collaborations with James Fearon and the Minorities at Risk dataset allowed him to do worldwide analysis of ethnic conflict, fully using formal theory, narrative, and quantitative analysis. This commitment to really learning new places (including languages) and new methods is extraordinary—and exceptional. The decision to take the time and effort to learn several languages and more than two methods is not something most scholars seem to want to do. Whether we agree or disagree with the merits of multi-method work, we have to acknowledge that beyond non-formal theory, one or two methods at most appears to be the norm in comparative politics and IR. As I have argued, resources, disciplinary norms, and the nature of the problems facing researchers all impact the methodological choices researchers make, but as the case of David Laitin shows, the will of the researcher may also be a factor. Even amongst social scientists, there's no accounting for taste.

Conclusions

Methodological choices cannot be dictated from without. Researchers must be free and encouraged to make choices appropriate to the problems at hand given the resources they can acquire. The above discussion suggests that while we should not expect to see all four methods being used by individuals in large numbers of scholarly works anytime soon, the barriers to collective diversity are not particularly high either and therefore we may well see all four methods more evenly represented in research problems, depending of course on the problems themselves, resources, disciplinary norms, and the preferences of individual scholars.

In addition Table 1, and the two distinctions of theory vs.

data, and formalization vs. non-formalization may suggest a way to bridge the idealism of the individual-based tripartite method with that of the more common one-or two-method scholarship. If researchers were to choose a method from each column and from each row, it would force most people out of their comfort zone. To fulfill this requirement, researchers would have to include one formal component (either formal theory or quantitative analysis) and one non-formal component (either theory or narrative); similarly they would have to include one data component (either narrative or quantitative analysis) and one theoretical component (either formal or non-formal theory). This type of selection rule in choosing methodologies would introduce much greater flexibility than Laitin's tripartite suggestion, by allowing for fewer methods in any one project and including non-formal theory as a choice. But, it would follow the spirit of Laitin's framework and his own work, by encouraging all researchers to bridge the mathematical and empirical divides. I hasten to add, however, that even this two-by-two framework and methodological selection scheme has to be seen as an ideal type predicated on the assumption of adequate resources, including data, theory, and skills of researchers. The most difficult and important problems that political science faces—e.g. democracy, development, and representation in inhospitable circumstances—may be areas where several types of resources are lacking, and therefore at the end of the day researchers have to make methodological choices given the demands and constraints of problems of interest.

Notes

¹ To be fair, his primary goal in the 2003 article is set up the tripartite method as a framework for social science and to place narrative within framework on equal footing with quantitative and formal modeling, and in the article Laitin spends a great deal of time discussing the value of narrative methods. So, it is possible that the vagueness is the unintentional result of a focus on other issues.

References

- Bates, Robert. 1998. "The International Coffee Organization: An International Institution," in *Analytic Narratives*, ed. Robert Bates et al. (Princeton: Princeton University Press), 194-230.
- Laitin, David. 2002. "Comparative Politics: The State of the Sub-discipline," in *Political Science: The State of the Discipline*, eds. Ira Katznelson and Helen V. Milner. (New York: Norton), 630-659.
- Laitin, David. 2003. "The Perestroikian Challenge to Social Science." *Politics and Society* 31:1 (March), 163-184.
- Przeworski, Adam, et al. 2000. *Democracy and Development: Political Institutions and Well-Being in the World 1950-1990*. Cambridge: Cambridge University Press.
- Stone, Randall. 2002. *Lending Credibility*. Princeton: Princeton University Press.

Mechanisms vs. Outcomes

Kanchan Chandra

New York University

kanchan.chandra@gmail.com

In his book *Hegemony and Culture*, David Laitin described himself as being committed to "a comparative politics that is sensitive to the particularities of each society, yet asks broad and general questions about all societies" (1986: xii). This idea of comparative politics—that it is in part a discipline that engages in the study of individual countries mainly for the purpose of producing cross-country generalizations—is the way in which most of us define the field now. And Laitin's work, which includes a study of the particularities of Somalia, Nigeria, India, Spain, Estonia, Latvia, Ukraine, and Kazakhstan in order to produce knowledge about other countries and continents, is unprecedented in comparative politics in its ambition and accomplishments in combining depth and breadth.

But what kind of breadth should we expect depth to generate? What kinds of generalizations based on within-country studies should we value in comparative politics?

In principle, we value generalizations about *outcomes*. So, when Lijphart finds that consociationalism preserves democratic stability in the Netherlands (Lijphart 1975, 1977), we want to know if consociationalism is also associated with democratic stability in other countries—South Africa or the former Yugoslavia. When Putnam finds that social capital explains institutional performance in Italy (Putnam 1993), we want to know if it also explains the same outcome elsewhere—Russia, or the U.S.. And when Laitin finds that the hegemony introduced by colonial rule explains the non-politicization of religion in Yorubaland (Laitin 1986), we want to know whether colonial hegemony explain the non-politicization of cleavages in other places—Zambia or India. Indeed, the ability to generate correct predictions about outcomes in out-of-sample countries is often treated as a test for the validity of a theory developed from a within-country study.

Against this backdrop, I make four arguments in this essay, illustrated with reference to Laitin's work:

(1) Although I share the view that the value of within-country studies in comparative politics lies in generating knowledge about other countries, I think that we are wrong in trying to distill generalizations about *outcomes* from within-country studies. The generalizations we should look for are generalizations about the *mechanisms* linking the independent and the dependent variable.

(2) We should evaluate the quality of such generalizations, not by testing to see if the entire chain of mechanisms linking the cause and the outcome in one country is the *same* in others, but by seeing how *far* the chain of mechanisms in a new country coincides with that of the first before it diverges.

(3) Arguing about whether we should use ethnography or rational choice or both in our work is beside the point. "Ethnography" and "rational choice" are not strictly comparable—the one is an approach to how data are collected, the other an

approach that tells us what to look for in the data. A forced comparison between the two suggests they are simply two overlapping ways of identifying mechanisms, among many other ways. They are not mutually exclusive—both can be used to identify mechanisms. Neither are they exhaustive. A scholar could use neither and still produce illuminating research. We should be arguing about what kind of knowledge to aim for, rather than privileging one or two approaches among a multitude of possibilities for generating that knowledge.

(4) The ambition of a mechanism-oriented approach is different from that of an outcome-oriented approach—and should produce different work. An outcome-oriented approach aims to use within-country studies to impose a uniform explanatory framework—with the same independent variables explaining the dependent variable—on as many countries as possible. The ambition of a mechanism-oriented approach is different: it is to use the study of one country to produce more sophisticated questions about new countries and get us closer to a series of unique point predictions about outcomes in other countries without the expectation that these outcomes will be the same in all countries.

Many of us in comparative politics would balk at this advocacy of point predictions. We tend to believe that “science” is about generalizing about outcomes across countries. I do not argue here that we should not aim for such generalizations in comparative politics. Some approaches that are, and should be, well accepted in comparative politics—such as cross-country statistical work—do indeed identify general patterns across countries that can be justified according to prevailing social scientific standards. But when working with within-country studies, there is a conflict between “science” and this type of generalization. When transplanting knowledge from within-country studies, it is scientifically more defensible to aim for unique point predictions in additional countries.

I read the progression of Laitin’s work in the past twenty years as an evolution from a concern with making generalizations about outcomes from within-country studies to a concern with making generalizations about mechanisms—and in particular, mechanisms about strategic action. Correspondingly, his work has also implicitly moved away from a concern with making predictions about outcomes intended to apply to a broad universe of cases to a concern with making a series of point predictions about individual countries. In the process, it has become more scientific.

I. Outcomes and Mechanisms

Outcomes and mechanisms have been distinguished in several ways. One way of making a distinction is to say that outcome-oriented analyses are about correlations (or laws), while mechanism-oriented analyses are about causal paths (Elster 1989, 3-10). But what is a causal path if not a series of correlations? Another is to say that the difference between outcome and mechanism is the difference between structure and process. But that just shifts the question—what is structure and what is process? A third is to say that outcomes are “tangible, observable things,” while mechanisms are unob-

servable mental processes. But what seems to be a “tangible, observable thing” is often a realization of unobservable idea—“democracy,” for instance, which we routinely think of as an outcome, is an intangible idea for which we have developed measures and indicators. And what seems to be an unobservable mental process is often something that can be observed and measured—for instance, we could measure “beliefs” by developing indicators based on survey data or experimental data that test for actions consistent with some beliefs and not others.

I take the difference between an outcome-based analysis and a mechanism-based analysis to be a difference in the *degree of the fineness* of the analysis rather than a difference in *what* is being explained. An outcome-based analysis makes a statement about a “macro-correlation” between two variables:

$$1 \rightarrow 2$$

Mechanism-based analyses are statements about the series of “micro-correlations” that constitute the logical chain linking the macro-correlation:

$$1 \rightarrow 1.1 \rightarrow 1.2 \rightarrow 1.3 \rightarrow 1.4 \rightarrow 1.5 \rightarrow \dots \rightarrow 2$$

I use a number series to illustrate the difference between the two because it illustrates nicely that there is no limit to the fineness of analysis that can be used in a mechanism-based approach. Just as there is an infinity of points between two numbers on a continuous scale, there is an infinity of micro-correlations that constitute the larger correlation.

But there is no difference in the essential nature of the dependent variable that either an outcome-oriented or a mechanism-oriented approach attempts to explain. The difference lies only in the distance of the explanatory variable from the dependent variable. In an outcome-based approach, the independent variable is distant from the dependent variable. The correlation between the two, therefore, is not obvious. In a mechanism-based approach, each micro-correlation links two proximate points. Thus, each micro-correlation may well be obvious. But a sequence of micro-correlations taken together travels a great distance from the initial variable that triggered them, and produces a non-obvious outcome.

II. Why We Should Not Expect Within-Country Studies to Produce Generalizations about Outcomes Across Countries

Within-country studies in comparative politics now routinely present themselves as using controlled comparison to identify hypotheses about outcomes that can be generalized to outcomes in other countries—or are evaluated on the ability to produce such generalizations. Laitin’s 1986 *Hegemony and Culture* is an example. The outcome that Laitin wants to explain in this book is the non-politicization of religion in Yorubaland in Nigeria. Although both tribe and religion are socially salient cleavages in Yorubaland, tribe is politicized and religion is not. The key independent variable explaining this outcome is the ideological hegemony instituted by the colonial state. British colonialism in Yorubaland adopted a system of indirect rule which created a commonsensical world in which tribe was “real” and religion was not. Consequently,

long after the departure of the British, the Yoruba organized their politics on the basis of tribe rather than religion. Within the context of Yorubaland, the simple bivariate correlation unearthed by the book is:

Colonial Hegemony →
Post-colonial Politicization of a Cleavage

If we wanted to use this within-country study to explain *outcomes* in another country what should we do? At first glance, it appears that we should look for the same simple bivariate correlation between the cleavage institutionalized by the colonial state and the cleavage that becomes politicized among the public. Indeed, this is what Laitin does in his concluding chapter, where he attempts to “test” the hypothesis developed from Yorubaland in the case of Benin. The book finds that the pattern of politicization in post-colonial politics can be explained by the pattern of politicization adopted by colonial rule. Consequently, Laitin argues that Benin “demonstrates the power...of the model of hegemony” (p. 165).

But if we really wanted to test the theory of outcomes in Yorubaland in other countries, the first step would be to note that this theory is not bivariate at all. It can be taken as a bivariate relationship *only in the context of Yorubaland*, when all other variables are held constant by virtue of the selection of the case. These control variables include any number of things such as the economy, the ethnic demography, the duration of colonial rule, climate, ecology, history, patterns of past violence, political leadership, and institutional structure, only some of which can be explicitly identified. These controls make it possible to isolate the impact of colonial rule in the case of Yorubaland. But when exporting the model to other cases, the researcher must also export the controls within which it is embedded, implicitly or explicitly. Thus, the test should really read:

Colonial Hegemony + Economy + Ecology + Climate +
History + → Post-colonial Politicization of
a Cleavage

Unless this entire model is exported, we do not know whether it is corroborated or not. The fact that Benin corroborates the model is thus a false positive, since we do not know whether the control variables, on which the effect of colonial hegemony was contingent in Yorubaland, also took on the same value in the case of Benin. And had Laitin reported a case which disproved the bivariate correlation—i.e., showed that the institutionalization of cleavages during colonial rule was not associated with the post-colonial politicization of cleavages—it would have been a false negative, since that simple correlation is not in fact the true model.

Could Laitin have tested the “correct” model? Could anyone, on the basis of a within-country study? I don’t think so. To conduct the right test would mean that all the control variables embedded in the case selection would have to be made explicit and entered into the “export” version of the model. But this is not possible, because it is not clear what those control variables are. There are an infinity of things that are controlled for implicitly when a researcher chooses a controlled design

within a country. That is what makes a within-country design attractive as a natural experiment. How might we put these many things into a model? A researcher might assert that not all of these controls are relevant—that only two of the possible infinity of variables matter and thus only two need to be exported to a new country or countries. But the research design does not permit that inference, since there is no variation in the control variables on the basis of which the researcher could have determined the relevance or irrelevance of these variables. And even if a model could be specified, we could not estimate it, because it would be heavily overdetermined, with more variables than cases.

I have used *Hegemony and Culture* as an example, but many within-country studies in comparative politics fall into the same mold. Like *Hegemony and Culture*, their titles advertise simple bivariate correlations. *Making Democracy Work: Civic Traditions in Modern Italy* (Putnam 1993) advertises a bivariate correlation between civic traditions and the performance of democratic institutions; *Ethnic Conflict and Civic Life: Hindus and Muslims in India* (Varshney 2002) advertises a bivariate correlation between civic life and ethnic conflict; *Institutions and Ethnic Politics in Africa* (Posner 2005) advertises a bivariate correlation between institutions and the politicization of cleavages; *Votes and Violence: Electoral Competition and Ethnic Riots in India* (Wilkinson 2004) advertises a correlation between electoral competition and ethnic riots; and my own work *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India* (Chandra 2004) advertises a bivariate correlation between patronage and the politics of ethnic headcounting. The correlations are surely contextualized by the names of countries and regions—“in Italy,” “in India,” “in Africa.” But the authors aim in these within-country studies to produce hypotheses that explain outcomes in other countries, and are evaluated on the basis of this ambition.

For the reasons given above, this ambition is unlikely to be realized, and this standard of evaluation unlikely to be met. Indeed, while we can think of numerous examples of within-country studies that identify interesting, internally consistent, bivariate hypotheses, it is difficult to think of bivariate hypotheses generated from within-country studies that have actually been verified through cross-national research.

III. Why Theories of Outcomes Within Countries Should Produce Generalizations about Mechanisms Across Countries

Had Laitin wanted to generalize about mechanisms rather than outcomes based on *Hegemony and Culture*, what might he have done? The first step would have been to conduct a far more finely-grained analysis. A disaggregated version of the analysis in the book as it stands is as follows:

Colonialism → Indirect Rule → “Religion expunged” from commonsensical assumptions about politics in the colonial period → Post-colonial Politicization of a Cleavage

The greatest level of disaggregation in the book is in the analysis of colonialism. But it jumps over the many links that

might conceivably tie colonial politics to post-colonial outcomes. A mechanism-oriented approach might have produced an analysis that looks as follows:

Colonialism → Indirect Rule → “Religion expunged” from commonsensical assumptions about politics in the colonial period → Additional Variable 1 → Additional Variable 2 → Additional Variable 3 → Additional Variable 4 →
..... → Post-colonial Politicization of a Cleavage

Generalizing from this study to a different country, then, would mean checking to see whether particular links in the chain of micro-correlations are replicated in another country – and, since we can almost be sure that it will not, identifying the point in the chain at which events in the other country diverge.

The purpose of such a generalization is not to “test” the hypothesis generated in the first country—a validation of that sort, I argued above, was not possible. But it is to lead the researcher to narrow the field of inquiry in the new country by taking her towards a narrower and more sophisticated set of questions.

Let me illustrate by showing how *Hegemony and Culture*, based on Nigeria, narrowed the field of inquiry for Posner’s inquiry into ethnic politics in Zambia (Posner 2005). Posner was interested in a similar question—what explains the pattern of politicization of cleavages in post-colonial Zambia? But Laitin’s book illuminated the place to start in looking for an explanation. Rather than having to start from the beginning and consider all possible hypotheses, Posner was able to narrow his field of inquiry to the effects of colonialism. He found that the first few links in the chain developed by Laitin in Nigeria, which showed how colonial rule narrowed the set of options open to elites in post-colonial politics, worked in the same way in Zambia—and then he filled in the remaining links from colonial to post-colonial politics through a new model of institutional politics that highlighted the role of electoral and party systems:

Colonialism → Indirect Rule → “Religion expunged” from commonsensical assumptions about politics → Party System → Electoral System..... → Post-colonial Politicization of a Cleavage

This analysis could of course be disaggregated still further. But this illustrates, I think, the contribution of Laitin’s study *Hegemony and Culture*, and the way in which it should be evaluated. It is not a study that yields generalizations about outcomes which validate the theory and elucidate universal (or even regional) patterns. I cannot think of any within-country study in comparative politics that does that. But its contribution lies in its identification of generalizations about mechanisms, which narrow the field of inquiry for others and produce more sophisticated point predictions.

Before going further, let me address two immediate questions:

(1) Why isn’t a mechanism-based approach subject to the same criticisms about controls as an outcome-based approach? If a macro-correlation cannot be exported to a new country

without its attendant controls, why is a micro-correlation exportable without those controls?

The critique of outcome-based approaches, which emphasized that the effect of one independent variable is contingent upon the presence of the control variables specified in the model, applies to enterprises which are trying to test a theory. If we interpreted the replication of each micro-correlation in a new country as a “test” of the argument produced by a previous within-country study, the same criticism would hold. But rather than thinking of the presence of each micro-correlation as a verification, we should think of it as a lever to identify whether or not there are variables that matter in the new country that did not in the first. If a micro-correlation is replicated, we should ask why, given the variables that characterize the new country. And when the paths begin to diverge, we should also ask why. Spotting repeated mechanisms in a new country need not mean that the same variables are at work in each country—the same mechanism might be produced by a different variable. By forcing the researcher to identify variables repeatedly in the course of building an explanation, a mechanism-approach is designed to uncover the importance of other variables—it may even be biased in favor of discovering additional variables.

(2) Why should we change our standards of evaluation from outcome-based approaches to mechanism-based approaches? Isn’t a mechanism based approach simply an imperfect rendering of an outcome-based approach? If so, we should continue to aim for generalizations about outcomes, since holding on to an ideal standard is important in inspiring research.

Our standards of evaluation determine the type of research we do. One reason to switch from standards that value generalizations about outcomes to standards that value generalizations about mechanisms is that this will lead us to look for different things. Had *Hegemony and Culture* been written in a field that valued mechanisms more or as much as outcomes, my guess is that the level of analysis would have been far more fine-grained than it was, and narrowed the field of inquiry for other researchers more than it did. Changing our standards may improve the quality of what we do in another way. Standards that value outcomes over mechanisms impose a direct contradiction between the scientific nature of the findings, which are contingent upon controls, and the demands of generalization in comparative political “science.” This contradiction normalizes an implicit disbelief in the way that research is done in comparative politics. Authors do not believe that they are producing generalizations about outcomes that apply across countries, but insert obligatory claims about such generalizations to be uncovered by “future research” in order to prove their professional credentials. Readers do not often believe these claims but look for them anyway, as the obligatory evidence of credentials by comparative political scientists. This institutionalized disbelief has a corrosive impact on the quality of research over the long run. Bringing our beliefs about what we should do more in line with what we actually can do may produce better quality research.

IV. The Movement to Mechanism-Based Approaches in Laitin's work

I read the evolution of Laitin's work since *Hegemony and Culture* over the last twenty years as a movement away from generalizing about outcomes to generalizing about mechanisms (Laitin 1989, 1992, 1998). The central question in this body of work is: what will the language outcome be in new states? Should we expect the development of a single national language? Laitin used a game theoretic representation of language outcomes in Western Europe to narrow the field of inquiry in India (Laitin 1989), the study of India to narrow the field of inquiry in Africa (Laitin 1992), and the studies of India and Africa, among other countries, to narrow the field of inquiry the former Soviet Union (Laitin 1998). The particular language outcomes in all these countries and regions were different. But, as I discuss below, the chain of mechanisms identified in each case helped to identify further mechanisms, and explain the particular outcomes, in every subsequent case.

The study of Western Europe generated chain of mechanisms that produced an outcome unique to Western Europe (Laitin 1992, 35-36):

Rationalizing policies of the state → Compliance by regional elites → Single language outcome

Starting with this initial path, Laitin identified a point of divergence in India:

Rationalizing policies of the state → Compliance by bureaucratic elites → Compliance by people (rather than regional elites) → Adoption of 3 ± 1 language outcome

The point of divergence in India from the path in Western Europe was driven by the timing of language rationalization, and the presence of different sets of actors. Language rationalization in Europe, which took place in a pre-democratic framework, was a matter decided upon by regional elites and those in control of the state. But the attempt at language rationalization in India took place after the introduction of democracy, and two other sets of actors were involved—bureaucratic elites, and the people, rather than elites, of individual regions. Laitin was able to map the path in India not by starting afresh, but by starting with the path identified in Western Europe and then identifying this point of divergence.

The path in India then illuminated a more complex path in African states, with several additional variables:

Rationalizing policies of the state → Compliance by bureaucratic elites → Compliance by people → Number and distribution of languages → Administrative Structure → Political Leadership → Unique Language outcomes (One language [Somalia], two languages [southern Africa], 3 ± 1 [possibly Nigeria] and so on)

As we see, the precise path varies from country to country. But it is not completely different—several mechanisms recur across all pairs of countries, and there is at least one mechanism that recurs in all countries, taken together. This recurrent mechanism is the “private subversion of a public

good.” Although individuals all cheer the adoption of a language rationalization policy on paper, it is not individually rational for them to switch to a single language themselves or educate their own children in it. Consequently, through a series of individual decisions they subvert the policy in practice.

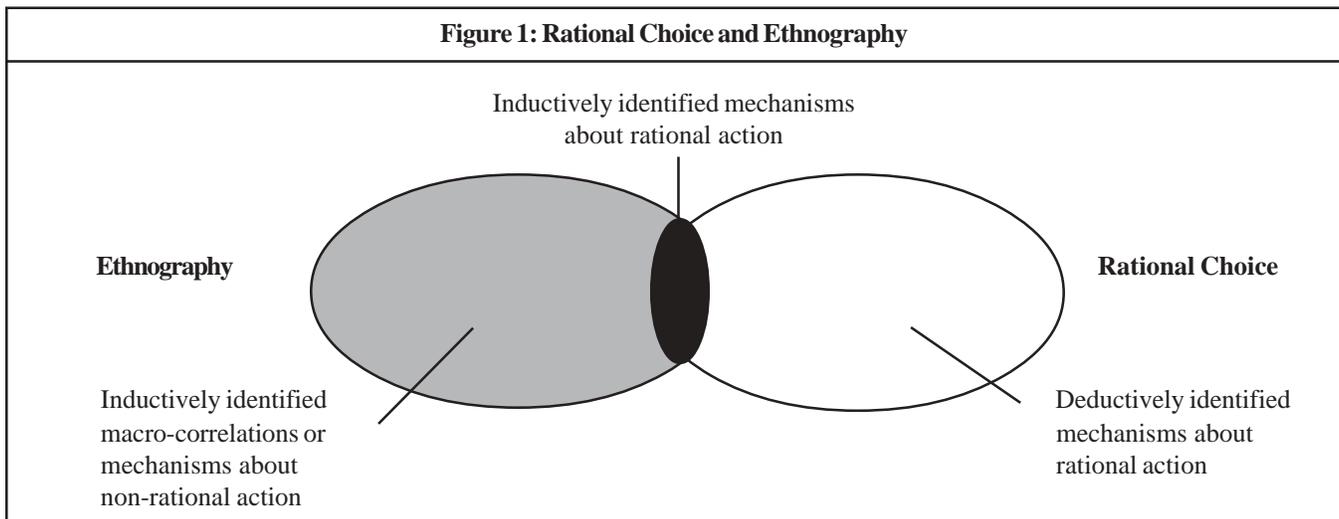
The result is a set of unique outcomes in each country, illuminated, but not determined by, the path of the previous countries studied. Indeed, Laitin notes: “The most important general finding of the game-theoretic analysis is that the ‘players’ involved in state construction are different over the centuries, leading to differently constituted language games and of different equilibrium outcomes” (Laitin 1992, 119). This is a very different type of analysis from *Hegemony and Culture*, where the existence of different equilibrium outcomes across countries would have been seen as a failure of the analysis rather than its most important finding.

In his most recent book *Identity in Formation*, Laitin takes just one link in the chain of mechanisms identified in previous work—the compliance of ordinary people with policies aimed at language rationalization—and disaggregates it still further, going down to an individual level of analysis which was missing from the first two books (Laitin 1998). This study, perhaps to a greater degree than any of the others, illustrates the extent to which adhering to the scientific standards of controlled comparison within a country makes it difficult to generalize about outcomes across countries. The book, which compares different linguistic strategies of Russian-speakers across the republics of Kazakhstan, Estonia, Latvia and Ukraine, makes good use of the strategy of controlled comparison to generate its findings—it focuses on the same linguistic group with the same history in the same region, in similar settings—and shows how variation in individual strategies is explained by a difference in the expected payoffs of compliance and the proportion of co-ethnics who comply across the four countries. The same controls that make this study scientific also make it difficult to generalize about outcomes across countries. Indeed, Laitin is careful not to make predictions about language assimilation outside the former Soviet Union. But what this study contributes to scholars studying other countries is an individual-level mechanism which may well recur across countries, and, when linked to other chains of mechanisms, explain unique outcomes in those countries.

V. Ethnography and Rational Choice are Two Overlapping and Non-Exhaustive Ways to Identify Mechanisms

“Ethnography” and “Rational Choice” are not strictly comparable. Ethnography is the firsthand personal study of a small group of people in a local cultural setting. As such it is simply an approach to how data should be collected, without assumptions about the kind of behavior the data should reveal. Rational choice, on the other hand, tells us how to model behavior, or how to interpret data on behavior, rather than telling us how to collect the data in the first place. The model of behavior requires self-interested, goal-oriented actors who choose between a finite set of alternatives based on a calculation of the

Figure 1: Rational Choice and Ethnography



costs and benefits.

To the extent that we can compare them anyway, there is an overlap between the two. Ethnographic methods can be used to probe the data for whatever relationships the researcher is interested in—the relationship between an independent variable and an outcome, or the sequence of mechanisms leading from the independent variable to the outcome. Ethnography might yield macro-correlations or mechanisms that are consistent with a rational choice model as well as macro-correlations or mechanisms that are not. Ethnography requires only that the macro-correlations or mechanisms are inductively generated in a particular way—by studying people in their cultural settings. A rational choice approach, on the other hand, is concerned explicitly with mechanisms that assume or reveal self-interested, calculated actions. These mechanisms may be deductively or inductively identified. Indeed, although it is common to associate rational choice theory with purely deductive thought, many of the most important rational choice arguments in comparative politics are derived from, and embedded in, particular empirical contexts—Down’s *Economic Theory of Democracy* (1957), for instance, is informed especially by the dynamics of the two-party democracy in the United States, and Riker’s theory of federalism rests on the particular path of American federalism (Riker 1964). The area of overlap between the two approaches thus lies in the area of inductively developed mechanisms about rational action. This overlap is represented in Figure 1 above

Each approach has its comparative advantages, allowing us to model certain types of mechanisms, at certain levels of analysis better than others. The problem dictates the choice of approaches or the combination of them. This is by now a standard pious statement in the methodological wars in comparative politics, usually used to argue that there is no one right way for how all research should be done. That’s true, but it also means that for particular problems, there *is* one right way, or a small set of right ways. Problems do dictate choices of methods, and if we do not adopt certain methodological approaches, we cannot solve certain problems.

Let me illustrate again through Laitin’s work. In his first book, *Politics, Language and Thought* (Laitin 1977), Laitin

argued that the choice of a single state language is not neutral—it creates (or maintains) an elite, composed of those who have access to education in the state language, and a subordinate class (or classes) of those who do not. Once we accept that the choice of official language puts some people in a position of disadvantage and others in a position of disadvantage, it is natural to ask how *expectations* about an outcome affect the outcome itself by affecting the strategy of political actors in the present. If English-speaking Somalis know that the choice of Italian rather than English will affect their job prospects in the future, then might this not affect the language that is actually adopted by informing their strategy in the present?

Rational choice approaches are the principal family of approaches which have been used to theorize about problems of this nature. Thus, if a scholar wants to model the relationship between expectations and outcomes, she must turn to rational choice methods. She may very well discover other approaches that are eventually superior, but this is the logical place to start. Not surprisingly, then, Laitin has tried to integrate rational choice into his study of language outcomes in every book on this subject in every book on this subject since the publication of *Politics, Language and Thought*. He has used ethnography, but also history, survey research and experimental methods to map the preference structures of individuals and collective actors—and then rational choice analysis to predict actions given this preference structure. Other research problems may well suggest different approaches to identifying mechanisms.

In recent work, Laitin proposes that this particular combination of approaches—formal theory, used mainly to mean rational choice theory, and narrative analysis, which I take to mean mainly ethnography and historiography—should become part of a standard, “tripartite” method in comparative politics, along with statistical methods (Laitin 2003a, 2003b). Given the context of the methodological debates in comparative politics, this appears to be a broad approach, suggesting, as he has demonstrated in his work, that both methods can be used in a complementary way and one need not be chosen at the expense of another. But if we think outside that context, this is an unreasonably narrow statement. Ethnography and rational

choice are not an exhaustive set of approaches which can be used to identify mechanisms. Depending on what kinds of questions the researcher is interested in, almost any technique can be used. Those that are being used in political science so far include survey research, experimental research in the laboratory or in the field, agent-based modeling, among others—and there are many other approaches that may be imported from other disciplines, including neurobiology, psychology, psychoanalysis, even literary analysis! Why restrict ourselves to only these two?

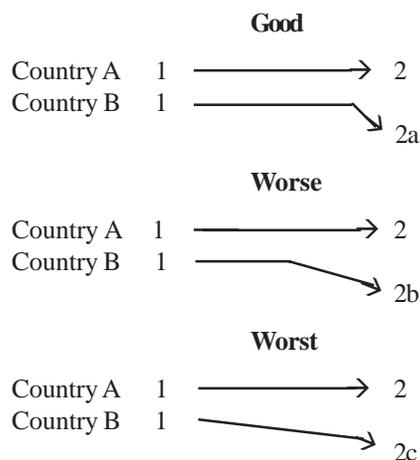
VI. How Should We Evaluate the Quality of a Mechanism-Based Approach?

The quality of a mechanism-based approach can be measured by the length of the chain of micro-correlations uncovered in one country that is replicated in a new country or countries. The longer the chain of micro-correlations in the new country that is illuminated by a previous study, the narrower the field of inquiry becomes in that country, and the more precise the remaining questions she needs to answer in order to generate a point prediction for that country.

Consider the possibilities represented in Figure 2. In each case, the study of Country A is an initial within-country study which produces a chain of mechanisms linking variable 1 to the outcome 2. We would like to generalize on the basis of this study to Country B. Note that in all three scenarios, the *same independent variable* produces *different outcomes* in Country A and Country B. But the three scenarios differ in the extent to which the chain of mechanisms triggered by variable 1 in Country B runs parallel to the chain of mechanisms triggered by variable 1 in Country A.

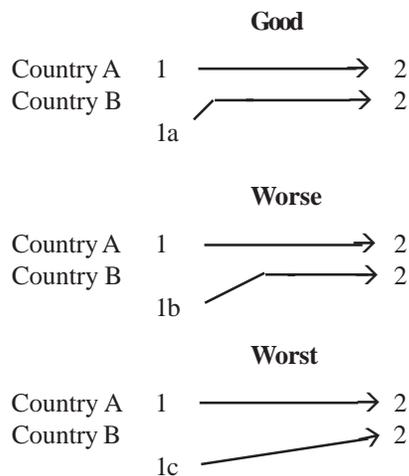
An outcome-based approach would judge all three scenarios to be of equally poor quality, since the same variable is not associated with the same outcome across both countries in any of the three scenarios. But a mechanism-based approach would judge them differently. In the first column, the chain of mechanisms identified in Country A dramatically narrows the scope of inquiry in Country B. In trying to explain the same outcome in Country B, we find that we can start at the same place as in Country A and follow along a great distance before the paths diverge. This is an example of a good generalization. In the second column, the chain of mechanisms identified in Country B takes us a smaller part of the way towards explaining an outcome in country B. This is an example of a worse generalization. And in the third column, identifying the chain of mechanisms in Country A that produce the outcome in Country A tells us nothing about the mechanisms that produce the outcome in Country B: the two paths diverge at the outset. This is the worst of the three generalizations.

Figure 2: Same independent variable, different outcomes across countries



We might also, in generalizing from one within-country study to another, move backwards from an outcome to a set of independent variables. Consider the scenarios below. In each scenario, the outcome is the same in Country A and Country B, but the initial *independent variable* that produces the same outcome is different. According to an outcome based approach, there is little to choose between the three scenarios—all three are bad. But the three scenarios are not equivalent according to a mechanism based approach. The first scenario shows a higher quality generalization about mechanisms than the second, since the chain of mechanisms replicated is longer. And the second scenarios show a higher quality generalization than the third, by the same logic.

Figure 3: Different independent variable, same outcome across countries



In each case, what the mechanisms identified through the study of one country can do in generalizing to other countries *is tell us where to look and what questions to ask—but not whether the answer is correct*. The test of whether each link in the chain of mechanisms that explain an outcome in a new country is the correct link should be a within-country test,

since only within-country tests can control for all the variables in that context. Continuing with the example of Laitin's work, if we want to test for whether the mechanism of the private subversion of a public good identified in Somalia recurs in India, we cannot assume that finding the same mechanism in both countries is evidence that it is at work in the same way across countries. We would want to identify observable implications of the argument about the working of the mechanism and test these implications using variation across space or time in India. It is through this painstaking series of questions and tests that we can get to unique point predictions for individual countries. And if the study of a phenomenon in one country shows scholars who study other countries which questions to ask, and which tests to perform, in order to generate point predictions for those countries, then that should count as progress in comparative politics regardless of whether these studies, taken together, add up to some universal explanatory framework.

References

- Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed: Patronage and Ethnic Headcounts in India*. Cambridge University Press.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.
- Elster, Jon. 1989. *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.
- Laitin, David. 1977. *Politics, Language and Thought: The Somali Experience*. Chicago: University of Chicago Press.
- Laitin, David. 1986. *Hegemony and Culture*. Chicago: University of Chicago Press.
- Laitin, David. 1989. "Language Policy and Political Strategy in India." *Policy Sciences* 22:4, 415-36.
- Laitin, David. 1992. *Language Repertoires and State Construction in Africa*. Cambridge: Cambridge University Press.
- Laitin, David. 1998. *Identity in Formation: The Russian Speaking Populations in the Near Abroad*. Ithaca: Cornell University Press.
- Laitin, David. 2003a. "The Perestroika Challenge to Social Science." *Politics and Society* 31:1, 163-94.
- Laitin, David. 2003b. "Comparative Politics: The State of the Subdiscipline." In Ira Katznelson and Helen V. Milner, eds. *Political Science: The State of the Discipline*. New York: Norton.
- Lijphart, Arend. 1975. *The Politics of Accommodation: Pluralism and Democracy in the Netherlands*, 2nd edition. Berkeley: University of California Press.
- Lijphart, Arend. 1977. *Democracy in Plural Societies*. New Haven: Yale University Press.
- Posner, Daniel. 2005. *Institutions and Ethnic Politics in Africa*. Cambridge: Cambridge University Press.
- Putnam, Robert. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.
- Riker, William. 1964. *Federalism: Origin, Operation, Significance*. Boston: Little Brown.
- Varshney, Ashutosh. 2002. *Ethnic Conflict and Civic Life: Hindus and Muslims in India*. New Haven: Yale University Press.
- Wilkinson, Steven. 2005. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. Cambridge: Cambridge University Press.

Understanding Rules and Institutions: Possibilities and Limits of Game Theory

David M. Woodruff

Harvard University

dmwoodr@fas.harvard.edu

Is rational choice theory compatible with, and useful to, ethnography, which I'll take to be the interpretation of meaningful action? For an affirmative answer, one might look to Fearon and Laitin's famous 1996 article on "ethnic peace." Their argument ran as follows: one way peace between two ethnic groups can be preserved is if each group punishes its own members for bad behavior toward the other group. Such "in-group policing" is effective, they suggest, because people usually have better intelligence about the doings of members of their own ethnic group. Thus, members of an ethnic group are in a position to reliably punish just those of their co-ethnics who have behaved badly in inter-ethnic interactions. By contrast, to the extent that people have a hard time identifying poorly behaved individuals who are not members of their ethnic group, they will only be capable of indiscriminate punishment of all the transgressors' co-ethnics. Such punishment may also deter bad behavior, but is more likely to lead to a spiral of violence. Fearon and Laitin capture these two possibilities in the form of two distinct equilibrium strategies in a repeated prisoner's dilemma game involving both intra-ethnic and inter-ethnic interactions.

To illustrate the real-world relevance of this argument, Fearon and Laitin provocatively mobilize an anecdote from the *locus classicus* of ethnography, Geertz's essay on "Thick Description."

Geertz relates that in early colonial Morocco, a marauding band of Berbers attacked the home of a Jewish trader in the Maghrib named Cohen. He survived but his guests were killed and his goods stolen. Cohen could get no help from the French authorities, but he belonged to a *mezrag*, or trade-pact system, and he went to his insurance broker, a tribal sheikh, to demand the assistance due. The sheikh knew precisely who had Cohen's merchandise, accompanied him in a climb up the Atlas directly to the shepherd of the thief's tribe, and took control of the entire herd. The tribal warriors soon returned, saw what had transpired, and prepared to attack. But then they saw Cohen and his insurance agent, a palaver began, and Cohen peacefully regained his goods at the precise insured value. [Cohen was given sheep meant to correspond to "four or five times" his loss (Geertz 1973, 8).] Note that 'on the equilibrium path' this institutional innovation of tribal 'information brokers' would make mutually beneficial trade relationships between Jews and Berbers possible, despite problems of opportunism due to a low density of social network relations. And, in the case Geertz relates, the institution also prevented spiral-

ing, here understood as a total breakdown of trading and relations between Jews and Berbers (Fearon and Laitin 1996, 728).

Thus, Fearon and Laitin see this as a clear example of the “in-group policing” mechanism; they imply that the sheikh’s action was motivated by incentives resembling those described in their model. However, let’s consider the other, extremely brief, account of the *mezrag* in operation provided in Geertz’s essay:

To make a trade pact in Morocco, you have to do certain things in certain ways (among others cut, while chanting Quranic Arabic, the throat of a lamb before the assembled, undeformed, adult male members of your tribe) and to be possessed of certain psychological characteristics (among others, a desire for distant things). But a trade pact is neither the throat cutting nor the desire, though it [the pact] is real enough, as seven kinsmen of our Marmusha sheikh discovered when, on an earlier occasion, they were executed by him following the theft of one mangy, essentially valueless sheepskin from Cohen (Geertz 1973, 12).

So in the two cases we see (1) the sheikh imposing material (sheep-denominated) damages for a large theft and two murders and (2) the sheikh executing seven people for a trivial theft. It seems plausible that in case (2) something other than the protection of Cohen’s property was at stake—perhaps the sheikh had contextual reasons of his own to show an iron fist to his tribesmen. And there’s no reason that something else couldn’t be at stake in case (1), as well—perhaps it was contextually important for the sheikh to intimidate the tribe in question (he did take a group of armed men with him when he went to capture the sheep). There is nowhere near enough evidence to sustain either of these interpretations—but there’s not enough to impugn them, either.

Any rational choice argument is built from two pieces: a vision of how people make choices, and a depiction of the circumstances under which those choices are made. The alternate interpretations of Geertz’s anecdotes I just presented rely on a presumption of rationally chosen action governed by (broadly understood) incentives (payoffs). With regard to how people make choices, the interpretations do not differ. They differ, instead, in their *depiction of the circumstances of choice*. While Fearon and Laitin relate incentives to repeated situations regulated by the trade pact, the plausible alternative interpretation describes incentives particular to individual situations where the rules of the trade pact happen to be in play. To use Weber’s terms, Fearon and Laitin focus on *formally* rational action, that is, on action motivated by bringing particular cases under general rules. The alternative contextual explanations describe *substantively* rational action, driven by a balance of considerations in a particular case.

I wish to make three points, focusing on understanding the intersection of rules (like those of the trade pact) and context. First, as just illustrated, the incentives that shape people’s actions in relation to a single rule can be different in different

situations. People can do the same thing—obey a rule, or enforce a rule, or violate a rule—for different reasons. The mere articulation of a rule that could allow formally rational action does not ensure that formal rationality governs in practice. My second point is that the potential relevance of case-specific incentives to rules has important empirical implications, affecting how we conceive the processes that allow institutions to create order. Third, even those who accept rational choice as a working assumption should conclude that game-theoretic methods of *describing situations* have little advantage when there are distinct contextual incentives for conformance to rules in distinct circumstances. Game theory is only powerful when the real-world incentives for conforming to rules are *general* and formal rationality governs. When substantive rationality looms large, game theory is at best not very powerful; at worst, the presumption that game theory will be powerful can obscure some of the key processes by which order gets built.

Contextual Incentives and “Cellular” Order

On reflection, it’s not very surprising that contextual, or case-specific, incentives can shape attitudes to rules. There are at least two ways this can happen. The first way is straightforward: the costs and benefits of conforming to or violating rules can vary across cases. Imagine Betty has contracted to buy sugar from Steve, with payment on delivery. While the sugar is still in transit, the price of sugar drops radically. How does Betty decide whether to pay or whether to weasel out of the contract? Let’s give a simple picture of her incentives:

| Action | Payoff |
|--|---------------------------------------|
| Fulfill contract | Sugar—contracted price |
| Weasel out of contract, buy sugar at lower price | Sugar—market price—costs of weaseling |

In other words, she pays when:

$$\text{market price} + \text{weaseling costs} < \text{contracted price}$$

The incentives in this case are contextual. The benefits of violating the rule (weaseling on the contract) depend on the difference between the contracted price and the market price, which may or may not be outweighed by the legal costs.

Actually, there’s no reason to stop with price swings. There can be an arbitrarily large number of contextual factors that determine the cost of weaseling to Betty (Commons 1957, 65). Does Betty have alternate suppliers? Does Steve have alternate customers? Are Betty and Steve linked in a kinship or religious network? Are there tax implications to paying or not paying? Does Steve’s business have enough working capital to wait out a lawsuit, or not? Is Betty’s firm under pressure from other creditors? All of these things will determine incentives relevant to Steve’s and Betty’s attitudes to the rule in any particular transaction.

In effect, Betty’s bill-paying behavior will reflect *multiple* equilibria in *multiple* games, in which the rules of contract are simply a single element. Some of these games will differ only

in the payoffs to different actions; others will feature moves, such as shifting suppliers, not always available. There can be an arbitrarily large number of such games, insofar as there is no guarantee that incentive structures and available moves repeat across transactions.

So far we've been considering cases where contextual incentives affect the costs and benefits of conforming to rules. A second, more complex form of interaction between rules and context occurs when a rule makes achievement of some contextual goal more or less expensive. It is important to remember that actions *consistent* with a rule are not always motivated by the general purposes behind the rule's creation. In other words, substantive rather than formal rationality can prompt the invocation of rules. Not long ago I was having lunch with another political scientist in Somerville, at a place that had tables out on the sidewalk. While we were eating, a man sat down at another table. He was middle-aged, light brown in color, and began muttering to himself as he sorted through the unrecognizable contents of two plastic supermarket bags. Within a few minutes the restaurant's cashier emerged to explain to him that tables were only for customers, and he apologized and left. Not much later, another person sat down at another nearby table—there were several available. This man was white, and looked old and tired. He did not mutter to himself or sort through his bags, but just sat for some minutes resting. The cashier did not emerge to run him off. It seems likely that there was some substantive difference between these two cases that motivated the invocation of the formally rational rule in one case but not in the other. In this light, it would be inappropriate to explain the invocation of the rule solely by the rule's terms. Similarly, when Geertz's sheikh had seven relatives killed as punishment for a trivial theft, the formal rules of the trade pact may well have provided a convenient occasion or pretext for a display of dominance needed for other purposes.

To summarize, then, a particular rule can intersect with contexts of varied incentives and opportunities. One potential contextual incentive is that the rule provides an economical means of achieving an aim—running off undesirables, intimidating challengers—that would have been sought even if the rule did not obtain. In these sorts of cases, game theory, I think we'd have to admit, isn't very powerful. All it amounts to is a kind of a protocol for writing down the results of ethnographic research in what could be a huge number of different models.

This isn't just a methodological point; it also has important implications for how one understands the way institutions create order. When incentives are contextual, order is *cellular* (cf. Stinchcombe 2001, 84, 97). General adherence to any rule reflects adherence to the rule in a multitude of "cells," each structured by distinct incentives and opportunities that interact with the rule in distinct ways. Things like life histories, relationships, and wealth—and, most importantly, any number of situation-specific goals that can't be brought under such general headings—matter greatly to the effectiveness of the rule. When incentives are contextual there can and often will be social situations where a rule is perfectly effective in guaranteeing order alongside others where the same rule, backed by the same state enforcement capacity,

does not lead to order. Let's go back to our sugar-purchase example. Suppose Betty notes that her competitor Bob tried to weasel out of a contract with Simone, but wound up losing a court case. Does she conclude that she has no alternative but to accept delivery of sugar at an above-market price? Well, it depends. What are Steve's incentives and capacity to wait out a lawsuit? How convincing is the technicality on which Betty plans to rely compared to that which Bob tried? And so on and so forth. Again, the same rule "fulfill contracts" may guarantee order for situations like Bob and Simone's but not like Betty and Steve's. Order obtains in some cells and not in others. Under cellular order, formally rational considerations are always potentially trumped by substantive rational ones.

General Incentives and "Broadcast" Order

It is not my intent to claim that actors encounter all rules in contexts that create idiosyncratic incentives. Some kinds of rules are associated with general incentives, which are the same any time an actor encounters a rule, whatever her other circumstances. Consider driving on a divided high-speed highway like America's interstates. Rules stipulate that on each side of the highway travel is only permitted in one direction. These rules are violated in a vanishingly small collection of cases. The incentive to travel in the right direction is that failure to do so involves an extremely high risk of a dangerous collision with oncoming drivers. This incentive is essentially general—not getting killed pretty much trumps any other considerations that one might have, such as getting to work faster, etc. Follow the rule, or else. Unlike the case of contextual incentives, with general incentives formal rationality and substantive rationality always both command the same action.

When incentives are general, game theory is powerful. It provides a compact way of expressing the structure of common choice situations. For instance, one can model the driving rule just discussed quite easily and convincingly as a coordination game. Again, the power of game theory here has empirical, not just methodological, implications. General incentives make order creation a "broadcast" phenomenon structuring the acts of general classes like drivers, rather than of individuals characterized by idiosyncratic incentives. Drivers' life histories, relationships, social status, wealth, desire to get to work on time, etc., are irrelevant. In effect, application of punishment in the form of crashing is entirely impersonal; so too is the reward of the much higher chance of reaching one's destination safely driving in the right direction.

There are cases of "broadcast" order-creation that are extremely relevant, even central, to comparative politics. For instance, there's clearly a difference between states in which organized violence by non-state actors must be conspiratorial and those where it need not be, and it would be both straightforward and convincing to give a game-theoretic account of this in terms of general incentives. (Unless there are a lot of publicly operating armed bands, any particular one will quickly be liquidated by a reasonably strong state, so getting to lots of such bands is a classic coordination problem that is hard to solve.) Likewise, some aspects of money

and property, capitalism's central institutions, rest on broadcast mechanisms of order. Or, to take an example from Professor Laitin's work, the structuring of choices on language can often be understood via coordination models that have a broadcast character. Nevertheless, rules surrounded by contextual incentives—affecting those subject to the rules or those enforcing them—are ubiquitous, and so too is cellular order.

Is Theoretical Modesty an Adequate Reaction to Contextual Incentives?

So far, I haven't challenged at all the idea that game theory could explain adherence to and violation of rules. All I've done is to argue that adherence to some rules can be modeled in a single game while adherence to other rules will require multiple games to explain. I've also argued that this line between compact and sprawling game-theoretical accounts also marks the border between broadcast and cellular order.

But isn't this dichotomy between "individual games for all the (many) contextual incentives relevant to a rule" and "one game describing general incentives to adhere to a rule" overly stark? Isn't there some middle ground on offer? After all, the task of social science would seem to involve transcending case-by-case storytelling (even storytelling employing the idiom of game theory).

I don't think the middle ground here is attractive. Cellular and broadcast order simply work too differently, meaning efforts to reconcile the sorts of contextual incentives characteristic of cellular order with the demands of tractability or compact modeling have high analytic costs. Consider the approach to contextual factors Fearon and Laitin take in their paper on ethnic peace. As they emphasize at several points, their "claim is not that the mechanisms we have identified are the only ones that matter, but that they have not been clearly identified before and do explain a part of the empirical puzzle" (Fearon and Laitin 1996, 727). Fearon and Laitin also point out specific aspects of context, such as state-building, that they have neglected.

Modesty of this sort greatly complicates straightforward verification of the argument with ethnographic evidence. Suppose we observe my hypothetical sugar-buyer Betty fulfilling a contract in a situation where the state has committed to punish her if she does not. Would someone offering a "partial" game-theoretic explanation of contract compliance as reflecting state enforcement be justified in concluding that Betty's action supports the argument? No! For it could be that the price differential was low, Steve was in a position to seek alternate customers, reputational costs were high, no more insistent creditors were on the scene, etc. Until all these other potential influences on Betty's decision were measured, her decision to fulfill the contract would say nothing whatsoever about whether enforcement explains "part of the empirical puzzle." Action merely consistent with the proposed partial explanation provides no support for that explanation, insofar as the action is consistent with a number of other plausible explanations. (Indeed, action inconsistent with the proposed partial explanation does not impugn it, either; it could

be that under other conditions enforcement would have been decisive in Betty's decision-making.)

Sorting out these issues requires a careful ethnography that seeks contextually relevant counterhypotheses for what was "off the equilibrium path," i.e., what other alternatives were open to Betty and why she did not choose them. The game-theoretic model might be good preparatory work for such an ethnography, insofar as it offers a plausible possibility for what was off the equilibrium path. But the claim of a partial explanation will collapse into a *de facto* negation of the importance of context if acts consistent with the model are taken as evidence of its operation. This temptation to misapprehend substantive rationality as formal rationality—to mistake cellular for broadcast order—is a significant analytic cost of offering game theory as merely a partial explanation.

Conclusion

I have argued that modeling of institutions as equilibrium strategies in a repeated game does work, in some circumstances—those characterized by general incentives and broadcast order. Here formal rationality and substantive rationality never work at cross purposes, meaning, among other things, that formal rationality cannot serve as an excuse for decisions made on substantive grounds. However, a danger arises when susceptibility of rules to compact game-theoretic modeling is assumed when incentives are in fact contextual. For if it is cellular order that a rule creates—if the rule determines behavior in some contexts and not in others, or if the rule can be mobilized as a convenient excuse—the correct game-theoretic depiction of the rule's operation would involve multiple games with distinct incentives and different available moves. Here the desire for a tractable model conflicts with the desire for an accurate one; when tractability wins out, the importance of context is negated and ethnographic precision becomes elusive. Modernization theory was criticized for arguments amounting to the suggestion that all substantive rationality empirically gives way to formal rationality (an argument, by the way, that would have been positively anathema to Weber!). Unless it takes context seriously, the game-theoretic account of institutions threatens to apply this same mistaken conclusion not just to "modernity," but to history far more generally.

One final point. The argument presented here has sought to demonstrate that there may be no easy way to transform a rule into a tractable game-theoretic model explaining the rule's effect on practice. But in many parts of the argument, one could easily replace "conformance to the rule" with any other explanandum of interest to political scientists—civil war, revolution, democracy, authoritarianism, tax incidence, etc. The serious difficulties facing a game-theoretic analysis of rules bespeak even more serious difficulties in applying the approach to these more complex explananda. For this reason, it would be a mistake to build comparative politics around the assumption that tractable game-theoretic models can provide satisfactory answers to our enduring questions.

References

- Commons, John Rogers. 1957. *Legal Foundations of Capitalism*. Madison: University of Wisconsin Press.
- Fearon, James D. and David D. Laitin. 1996. "Explaining Inter-ethnic Cooperation." *American Political Science Review* 90:4, 715-735.
- Geertz, Clifford. 1973. "Thick Description: Toward an Interpretive Theory of Culture." In *The Interpretation of Cultures: Selected Essays* (New York: Basic Books), 1-30.
- Stinchcombe, Arthur L. 2001. *When Formality Works: Authority and Abstraction in Law and Organizations*. Chicago: University of Chicago Press.

Ethnography and Rational Choice in David Laitin: From Equality to Subordination to Absence

Ted Hopf
Ohio State University
hopf.2@osu.edu

When we try to interpret politics in Africa (or anywhere, of course!) in terms of our own structures of preference and categories of action, we learn less about either Africans or ourselves than we do by recognizing that our political understanding is not universal, but is contingent on our sociological and historical experience (Laitin, 1986: ix).

In these opening passages, David Laitin rejects the false promise of positivist imposition of apriori categories on evidence, and embraces the interpretivist approach of inter-subjective contextualization and observational reflectivism. In other words, he lets his subjects speak; their versions of reality are the versions that matter, not his. It is intersubjective reality that matters, the web of meanings shared by a community, not any objective reality, or what is there independent of anyone's perceptions of it?

In the three works whose methods and methodological consequences I describe here, the common problem is political mobilization around multiple identities. In each work Laitin explains why some identifications, and not others, are fertile ground for political action.

In *Hegemony and Culture* (*HC*), Laitin stages a dialogue between interpretivist evidence and rational choice models, allowing the latter to frame what general conclusions he draws from his cases, but allowing his ethnography to govern the substantive content of these theories, modifying them in the process. This is sometimes called abduction, the conversation between theory and evidence, modifying both.

In *Identity in Formation* (*IF*), Laitin fixes a different relationship between ethnography and apriori models. Ethnography, the co-star in *HC*, is reduced to a supporting role in *IF*. If tests generated by the apriori model, in the form of surveys, experiments, and the statistical analysis of both, can be sup-

ported by Laitin's ethnographic recovery of the intersubjective worlds of his subjects, then this evidence is advanced in support of the hypotheses derived from the apriori theoretical frame. But virtually never does the evidentiary power flow in the opposite direction, as it did in *HC*. Laitin's ethnographic evidence in *IF* never modifies his apriori theories; it merely appears when it is supportive of those models.

In "Explaining Interethnic Cooperation," (EIC) written with James Fearon, intersubjective reality disappears as an object of theoretical or empirical interest. All social phenomena are either objectively labeled, or assumed to have a particular value. This article received an APSA prize as the best article to appear in the 1996 *APSR*. Clearly, objectivism has its rewards.

In *HC* Laitin demonstrates that rational choice and ethnography are not necessarily antithetical to each other, or even incompatible. Indeed, they may be most fruitfully combined.¹ In *IF*, however, the value of ethnography is barely more than supplemental to the evidence gathered through the more objectivist means of survey research and experimentation.² In EIC, evidence, let alone ethnographic recovery of intersubjective reality, is absent altogether.

Hegemony and Culture: Ethnography and Rational Choice as Co-Stars

In *HC* Laitin explains why the Yoruba in Nigeria are politically mobilized by some identities, but not others. His ethnographic findings had several counterintuitive turns. The default explanation going in was that religious identities were evoked by political entrepreneurs. But Laitin finds that tribal identification with an ancestral city, rather than with Islam, was the axis of political mobilization. And the local subjects offered prima facie evidence for that hypothesis, denying any religious differences between Muslims and Christians. Laitin's ethnographic research, however, demonstrated that while his subjects said one thing, they practiced another. Their mundane daily practices clearly showed differences between Muslims and Christians, so different religious identities, despite denials of relevance from subjects themselves, existed, and were enacted (Laitin 1986, 55-75). In interviews, however, Laitin found that these religious differences did not correspond to differences in the political views of his subjects. Digging still deeper, he finds out why this disconnect occurs between subjects' perceptions, practices, and political actions. They understand religion as something each chooses, while identification with an ancestral city is primordial and naturalized.

As Laitin wrote, while we can assume goal-oriented behavior, "only a theory of culture can tell us what goals are being pursued" (Laitin 1986, x, 11, 16, 104-5). He saw ethnography as a way of "adducing cultural preferences without tautologically claiming that preferences can be derived from the behavior of actors who are assumed to be rational" (Laitin 1986, 16). Had he simply made rational choice assumptions, he never would have been able to understand how, despite the material decline in importance of ancestral cities, Yoruba identification with these places, did not decline. "Rational choice theory cannot adequately adduce differential prefer-

ence functions across cultures” (Laitin, 1986, 179). Ethnography is necessary to find data on preferences, though once “within a cultural framework, it is useful to use rational choice” (Laitin, 1986, 182).

Identity in Formation:

Ethnography in a Supporting Role for Rational Choice

In *IF* Laitin explains the variance in the political mobilization of Russian-speakers in the post-Soviet space. His primary finding is presented as a tipping model, wherein he specifies various material goals for local Russian-speakers, their choice to assimilate or not to the titular nationality being determined in strategic interaction with other Russian-speakers.

Ethnography and rational choice are equal partners in *HC*, each informing and adding to the other. In *IF*, on the other hand, rational choice is the senior partner, ethnography the summer intern.

Laitin argues that his ethnographies are “representative,” presumably meaning worthy of evidentiary value, because “they are in conjunction with large surveys” (Laitin 1998, x). Laitin made no such superfluous extra justification in *HC*. But the results of these surveys come to dominate as evidence, his ethnographic data being advanced only if they support findings made in the surveys. I could not find a single case where ethnographic data were advanced as evidence in contradiction to the survey data, or still less, as evidence to interpret the survey data that were gathered. The intersubjective world of post-Soviet subjects was accorded far less evidentiary value than the answers to survey questionnaires, questions which were developed in light of apriori theories of the researcher, not from the ethnographic materials he gathered. This method is directly opposed to the epigram with which I started this article, the selection from *HC*.

For example, Laitin concludes from his survey data that Russians living in Latvia see little economic point in learning the Latvian language. But in a footnote, he reports that this finding “did not come through in our ethnographic fieldwork” (Laitin 1998, 117, n14). What is striking is that contradiction required no commentary, manifesting the very low regard Laitin accorded his own ethnographic evidence, relative to survey data. It would be hard to imagine this occurring in *HC*.

While daily social practices were a critical piece of evidence in *HC*, being both explicitly counterposed to the results of subject interviews, and accorded greater evidentiary weight than the latter, for example, in determining just how important religious identities were, in *IF*, they drop out of the story, and their absence exerts a powerful theoretical effect on Laitin’s conclusions.

One conclusion in *IF* is that “Russian-speaker” is a more important identity in the post-Soviet world than “Soviet.” But I would argue this finding rests in large part on an over-reliance on apriori theory and the answers to survey questions generated as hypotheses to test that theory. Had Laitin systematically evaluated his ethnographic data, he might have found that the Soviet identity was a set of lived daily practices, if not an answer to a survey question.

But even the survey data do not conclusively support

Laitin’s conclusions about a Russian-speaking identity. He reports that in the waning days of the Soviet Union, over 70% of Russians living outside of Russia claimed the Soviet Union, not Russia, as their homeland, or *rodina* (Laitin 1998, 308). Perhaps had the Soviet identity been measured differently, ethnographically, one would have found that the lack of Russian ethno-national mobilization against titular nations was consistent with an omitted variable: Soviet identity. In understanding themselves as multinational, as Soviet, the ethno-national axis of identification was simply not as available, or salient, to millions of Russians living abroad. Therefore, they did not understand themselves in opposition to Kazakhs, Ukrainians, and Estonians, in precisely that way, just as the Yorubas did not mobilize along Christian and Muslim axes of identification.

Daily practices in *IF* are reduced to language choice. Despite repeatedly, and correctly from an ethnographic point of view, stating that language is insufficient to measure identity, to capture the broad array of daily social practices that constitute identity, Laitin ends up relying on what he has already ascribed as inadequate: language. And he does so because of his sacrifice of ethnography to surveys, his desire to easily manipulate data, rather than wrestle with the likely complexity an ethnographic account of Russian intersubjectivity would offer. Laitin frankly states that he uses language because it is easier than culture, easier to measure and easier to monitor (Laitin 1998, 368).

Although Laitin acknowledges that language adoption does not equal identity adoption, in the body of the study, it is reduced precisely so (Laitin 1998, 23). Laitin defines assimilation as the adoption of the “cultural practices” of the other, but these data are never advanced to make the case, one way or the other (Laitin 1998, 152). He even offers a list: dress, manners, religion, food, artistic taste, slang, music, holidays, etc. but this kind of ethnographic data is not offered in the book (Laitin 1998, 297, 368). In the ethnographic protocol Laitin provides the reader (Laitin 1998, 394-5), we find that language use at home and work, reading habits, content of family library, radio and TV listening and viewing patterns, newspapers read, recreational activities—all rich sources of easily measurable ethnographic data, have been gathered. But these are never presented in the book, again implying that they were not used to specify the theory in an abductive way, to guide interpretation of survey and experimental results, or still less to reconstruct the intersubjective world of Laitin’s subjects.

Unlike in *HC*, where a rational choice model was informed by ethnographic evidence, Laitin allows the model here to drive the search and interpretation of evidence. Russians abroad decide to assimilate, or not, based on the relative costs of alternative strategies. This would be consistent with any statement of preferences, including those in *HC*. But Russians’ understandings of the costs are not ethnographically recovered, but instead assumed, or derived from survey questions whose formulations are already derivative of the theory being tested, not generated from the intersubjective understandings of the Russians themselves. For example, one might

hypothesize that costs vary according to Russian identification with the titular nationality, such that already understanding oneself as a Russian-Estonian reduces the apriori costs of assimilation. One could also hypothesize that a Soviet identity also reduces the apriori costs of assimilation because of its partial transcendence of ethnonational axes of identification. Such identifications would be recoverable ethnographically, through the array of practices Laitin already identified.

If we turn to the tipping model, we can see the consequences of ignoring ethnographic recovery of preferences, again something which Laitin defended as necessary in *HC* (Laitin 1986, 21-9). How do we know how many Russians have to assimilate to the local nationality before I myself will choose assimilation? Laitin offers several plausible accounts, but are they the ones Russians themselves used? This was not part of Laitin's research; instead they were taken as given, as if all Russians, or at least a theoretically meaningful number of them would calculate as assumed (Laitin 1986, 248). But this not only ignores the value of ethnography, but applies Schelling's model in a way Schelling himself explicitly guarded against.

Schelling's key variable was the number of houses in a neighborhood of eight you could tolerate being of a different race. He demonstrated thereby that even those who wanted integration, say a 40% minority, could still end up in a completely segregated neighborhood, if any one of their eight neighbors had preferences, and neighbors, that raised the minority representation to four houses. Schelling systematically varied the value of this factor to induce tipping (white or black flight) under different conditions. But Laitin does not hypothesize different Russian cost-benefit calculations, specifying the range in which to expect tipping to assimilation. He instead specifies that which Schelling varied for heuristic purposes, and Laitin in *HC* recovered through ethnographic research.

“Explaining Interethnic Cooperation”: Ethnography? What Ethnography?

To the extent that ethnography is the discovery of the intersubjective world of a community of interest, Laitin's article with James Fearon is the very antithesis of ethnography. Not only is no ethnographic research executed, but no evidence of the views of subjects of any kind is considered necessary to produce an explanation of why ethnic groups, despite incentives against cooperation, do so in any case.

Instead, the article develops a possible theory through a string of assumptions. It should be said that if each of these assumptions were true, the broad conclusions of the article would most likely follow. But how are we to know whether any of the assumptions are at all likely to appear in the world, absent any empirical, let alone, ethnographic, research?

It is striking that the kinds of meanings so painstakingly reconstructed by Laitin in *HC* through ethnography and in *IF* through survey research and experimentation are simply assumed to have one presumed value in *EIC*.

For example, they assume that the mechanism for defeating opportunism, self-interested behavior that has socially

harmful consequences, is reputation in a relationship. And this reputation for cooperation and trust is maintained through expectation of iteration (Laitin 1996, 717-8). So far, so PD. But why assume the need for iteration? Aren't there many identity relationships that are sustained just by being a particular identity? Isn't there a huge literature on group identity, in experimental psychology, cultural sociology and anthropology, and ethnography, that argues, and empirically demonstrates, that sharing (racial, ethnic, religious, sexual, gender) identities in and of themselves increases trust? Isn't the assumption of a need for iteration here a move dictated for model specification, rather than one grounded in consensual empirical research? And isn't it precisely the opposite of what Laitin did in *HC*?

A list of contestable assumptions follows. Each is either logically or empirically questionable, or both. And each appears consistent with the needs of an apriori theory, rather than reflective of the warnings in the opening Laitin epigram. It is revealing that virtually all the assumptions made are derived from game theory, not theories of identity, let alone from an ethnographic recovery of intersubjective reality.

They assume that ethnic groups are characterized by relatively dense social networks and low cost information about the past history of an individual's behavior. This assumption would appear to be necessary for the iteration assumption to operate. But what evidence is there that ethnic groups are particularly characterized as Laitin and Fearon posit? What if there are other social networks, equally or more dense, in which an individual participates? Laitin and Fearon note that interaction with a variety of groups dilutes the capacity to form reputations. Is this perhaps why they assume such interaction away, instead stipulating that an individual only belongs to one group (Laitin 1996, 719) !?

This stripped-down identity assumption then allows additional assumptions to be made about the greater information one has about one's own ethnic group members' conduct, the capacity to punish that behavior, the inability to punish behavior of someone in another ethnic group, and about each individual having only one identity, an ethnic one. Allowing additional identities, say religion, class, locality, or gender, would seem to complicate the apriori theory too much, as they might cross-cut or reinforce cooperative or hostile relationships instantiated by the ethnic identity relations (acknowledged in a small section in Laitin 1996, 727). Where has it been demonstrated that such a barren field of identity has ever existed in any reality on earth? This complete ignorance of the Other, except for her binarized ethnicity, then feeds into the game theoretic model, such that one knows nothing about the Other except her move to defect or cooperate. This yields predictable outcomes, but where? Where are the necessary conditions specified in this model to be found anywhere in the world past, present, or future?

In concluding, the authors make a rather jarring claim, given the above: “Our empirical section illustrated the plausibility of our assumptions....” I must confess I never found that section, and certainly there is nothing in the article that would qualify as empirics in either *HC* or *IF*.

Concluding Remarks

In this essay, I merely wanted to describe Laitin's work in terms of ethnography and rational choice, not evaluate it. The differences are clear across the three pieces. It is also obvious that each piece produces different kinds of knowledge, especially with respect to the expected generalizability of the findings. In *HC*, a rationalist theory of identity has been married to a particular set of empirical circumstances. In *IF*, a more elaborate rationalist theory of identity has been tested in a much larger, though still bounded, domain. In *EIC*, a universalist theory of identity has been tested nowhere, but has been demonstrated valid within a set of ultra-constraining assumptions.

Is the obvious true? Is ethnography the enemy of generalization? Perhaps in practice, but not in principle. Wedeen (2002) has recently written about the possibility of collecting intersubjective data based on phenomena such as identity, so that conceptualization of variables need not be derived exclusively from a priori theories, but rather can remain true to the ways in which concepts are understood in context. This would provide more reliable and valid data for those with statistical inclinations, for those who wish to specify survey and focus group instruments, and for those who wish to construct models with grounding in some reality.

The objective should be to return to Laitin's original insight. Theories of political action, of identity, of mobilization and identification require accounts of preferences that are not merely assigned, but theorized and empirically uncovered. And preferences themselves are not just the oral statements or written testimonies of subjects, but are embodied in their mundane daily practices. Ethnography, in this sense, is necessary for rational choice to produce creditable knowledge claims of any kind.

Notes

¹ Reminiscent of the efforts made to do so in Robert H. Bates, Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry R. Weingast, *Analytic Narratives* (Princeton: Princeton University Press 1998).

² This is reminiscent of Achen and Snidal's recommendation that qualitative case studies are best suited as the raw material appendages of rational choice models. Christopher H. Achen and Duncan Snidal, "Rational Deterrence Theory and Comparative Case Studies," *World Politics* 41:2 (January 1989), 167-69.

References

- Fearon, James D. and David D. Laitin 1996. "Explaining Interethnic Cooperation," *American Political Science Review* 90:4 (December), 715-35.
- Laitin, David D. 1986. *Hegemony and Culture. Politics and Religious Change among the Yoruba*. Chicago: University of Chicago Press.
- Laitin, David D. 1998. *Identity in Formation. The Russian-Speaking Populations in the Near Abroad*. Ithaca: Cornell University Press.
- Wedeen, Lisa. 2002. "Conceptualizing Culture: Possibilities for Political Science," *American Political Science Review* 96:4 (December), 713-28.

Recognizing the Tradeoffs We Make

Ashutosh Varshney¹

University of Michigan, Ann Arbor
varshney@umich.edu

David Laitin defies a famous binary classification of scholars between hedgehogs and foxes. The late Isaiah Berlin's work, following Tolstoy's, gave this distinction considerable currency in the social sciences. The hedgehog knows one thing very well; and the fox knows quite a few things, if not each in great detail.² Hedgehogs work on one given topic/theme/theory for an entire lifetime, adopting a cumulative research program, attempting to resolve one puzzle at a time, as they advance. Think of Arend Lijphart's lifelong pursuit of the idea of consociational democracy.

Foxes move from one big topic/theme/theory to another, each topic keeping them engaged for a few years but not more, showing enormous intellectual breadth in the process. Consider Samuel Huntington in political science, and Amartya Sen in political economy. Huntington has provoked new debates in three fields of our profession: comparative politics, American politics, and international politics. Sen ranges from rationality on one hand to famines and poverty, inequality, choice of techniques in planning, and, increasingly, identity on the other.

Laitin has worked almost entirely on ethnic politics, rarely if ever on development, economic reforms, democracy and authoritarianism, party politics, etc., let alone in other subfields of the political science discipline. Yet three things separate his work from a classic hedgehog strategy. His substantive questions have varied, even if the subfield has not; he has moved from country to country in search of answers; and what is most pertinent to this symposium, his methodological commitments have radically changed over time.

Three of Laitin's books deal with language politics. In *Politics, Language and Thought: The Somali Experience* (1977), Laitin probed the political and social consequences of maintaining a neocolonial language like English, as opposed to using a vernacular like Somali, as an official language. In *Language Repertoire and State Construction in Africa* (1992), he explained how very few African states went for linguistic rationalization in the classical European sense of having only one language, but many others went for two other linguistic strategies: a 2-language outcome, and what he came to call a 3+1 solution, a formula he found in India and has applied to other countries as well. In *Identity in Formation: The Russian-Speaking Populations in the Near Abroad* (1998)—*IF* hereafter—the central issue is how to explain the emergence of a "conglomerate identity," based primarily on linguistic adaptation, among the Russian-speaking populations of Estonia, Latvia, Ukraine, and Kazakhstan after the breakup of the Soviet Union.

His work on identity politics is, of course, not entirely driven by language issues. In his second book, *Hegemony and Culture: Politics and Religious Change among the*

Yorubas (1986)—*HC* hereafter—Laitin asked why Yorubaland's religious life was split between Muslims and Christians, but Muslim-Christian differences were not the principal cleavage in Yoruba politics. Finally, in two co-written articles with James Fearon (Fearon and Laitin 1995, 2003), he has examined the consequences of ethnic diversity for peace and violence. In the first joint article, Fearon and Laitin probed the conditions under which ethnic diversity would actually lead to peace, not violence; and in the second article, they asked whether ethnic diversity was indeed a crucial determinant of civil wars, or whether other factors were more significant.

The range of these questions makes Laitin a formidable intellectual force, indeed a central figure, in the subfield of ethnic politics. One can no longer write about language politics, identity formation, or ethnic peace and violence without engaging his arguments. Moreover, his frequent forays into new empirical terrains add greatly to his output. His case materials have come from Somalia, Nigeria, India, Sri Lanka, Catalonia, the Baltic Republics, Central Asia, and Ukraine. As my own research has become multi-country, it is now clear to me that developing intimacy with new political and cultural materials, a prerequisite for thoughtful work, is not easy. Consequently, I have developed a strong admiration for those who step beyond the existing zones of familiarity and develop ideas on that basis. Laitin has repeatedly allowed his intellectual curiosity to migrate to newer lands, also sometimes linguistically retooling himself. Many have worked on multiple countries; very few have learned new languages. Laitin may have never left the subfield of ethnic politics, but his intellectual journeys within have a fox-like quality.

Laitin's substantive achievements are not the principal issues for this symposium. Rather, our focus is on his methodological moves. The symposium seeks to assess the value of Laitin's methodological voyage from his early work based primarily on ethnography, whether in Somalia or in Ile Ife in Yorubaland, Nigeria, to his work over the last decade and a half, in which a rational choice stance has played a major role.

But before I proceed further, one should note that ethnography and rational choice are not the only methodological alternatives which should be discussed here. Rational choice methodology, which does tend to rely heavily on formal logic and a priori assumptions, as opposed to ethnography which is more empirically driven, is only one of the elements in Laitin's transformation. Some of his more recent work is heavily statistical, and we must draw a distinction between formal and statistical reasoning. If in "Explaining Interethnic Cooperation" (1995; *EIC* hereafter), assumptions, formal reasoning, and equations abound, and only illustrations from the real world are given but no systematic empirical evidence, in "Ethnicity, Insurgency and Civil War" (2003; *EICW* hereafter), there are only a few basic assumptions made and recourse to formal logic is minimal. Instead, existing theories of civil war are tested against a large statistical dataset. This kind of work is not ethnographic but it is still empirical, to be differentiated from the rational choice tilt of the *EIC*.

So how should we judge Laitin's methodological transformation? My central argument in this essay is that methods

always entail a tradeoff. Each method in the social sciences can handle some puzzles better, leaving others unresolved. It is not as helpful to say that ethnography is better than rational choice, as to determine what their respective strengths are, and what they can handle best. Creative imagination can allow us to blunt the edges of the tradeoff, but the tradeoffs do not altogether disappear.

In what follows, I will discuss this idea concerning each of the methods Laitin has deployed: ethnography, surveys, and formal modeling. I will primarily use his work as an illustration of tradeoffs, though in the process I will also discuss other work. All three methods have their unique mix of strengths and weaknesses, and we need to decipher what kinds of questions can be best analyzed by each. It is both a question about the strategy and substance of research. Laitin has not always been conscious of this point, nor has he consistently followed it. Since he subscribes to the notion of cumulation in social science, his recent critiques on purely methodological grounds seem quite puzzling and paradoxical. Basically, the form his critique has taken and his commitment to the idea of cumulation are not logically consistent.

Shifts of Evidence, Shifts of Method

Theoretical shifts, especially in light of changing evidence, are quite common in scholarly life. Robert Dahl became skeptical about the pluralist nature of American democracy, once the tight hold of business over American politics became clear to him in the 1970s (Dahl 1982). And as the revolutions overthrowing Communism squarely questioned his assumptions about human behavior, Jon Elster developed serious self-doubt about rational choice theories in the early 1990s (Elster 2000).

These are examples of evidence-based theoretical shifts. Are Laitin's shifts evidence-driven, or method-driven? In his thoughtful essay for this symposium, Ted Hopf (2005) suggests that the reasons are methodological.

But are methodological shifts entirely uncommon? In one of the famous interpretations of Marx's overall body of work, Louis Althusser argued that there was an "epistemological break" in Marx after his early work—before *Das Kapital* was written (Althusser 1969). According to Althusser, "early Marx" moved from the pseudo-scientific methods, when he gave too much emphasis to human consciousness, to science later when he made it unambiguously clear that the structure of production determined the relations of production and, therefore, human consciousness. Epistemology is about the ways of generating knowledge. Willy nilly, it becomes inescapably methodological.

In short, method-based theoretical turns have precedence in intellectual history. Like Marx in his later works, Laitin today tends to start with some universalist, a priori assumptions. Laitin, of course, does not leave it there, and unlike so many rational choice scholars, he does field work as well. But given that his survey questions are based on a priori theory, says Hopf, his empirical testing has become partial:

I could not find a single case where ethnographic data

were advanced as evidence in contradiction to the survey data, or still less, as evidence to interpret the survey data that were gathered. The intersubjective world of post-Soviet subjects was accorded far less evidentiary value than the answers to survey questionnaires, questions which were developed in light of a priori theories of the researcher, not from the ethnographic materials he gathered (Hopf 2006, 18).

This way of generating knowledge, Hopf continues, is “directly opposed” to the celebrated opening lines of *Hegemony and Culture* (*HC*):

When we try to interpret politics in Africa (or anywhere, of course!) in terms of our own structures of preference and categories of action, we learn less about either Africans or ourselves than we do by recognizing that our political understanding is not universal, but is contingent on our sociological and historical experiences.” (*HC*, ix).

Hopf is insightfully identifying the difference between surveys and ethnographies here. Though scholars select their ethnographic sites for theoretical reasons, surveys are more theory laden than ethnography. Survey questions are theoretically framed: only some questions are asked, not all possible questions. Ethnography facilitates a much more open-ended “soaking and poking” and, as Hopf puts it, “it lets the subject speak.”

Hopf is right about this, but it is also worth asking whether ethnography has some limitations and surveys some advantages. Ethnography clearly allows us to deepen, but surveys make broadening possible. Deepening and broadening as categories of empirical observation generate trade-offs. Ethnography makes accuracy about a case or two possible in a way that surveys can not match; but surveys allow a broader range of observation, covering many more cases than ethnography can possibly do. I find Laitin’s belief in *Identity in Formation* (*IF*) that ethnography alone would not take him forward in the Near Abroad well founded.

Though *HC* and *IF* seem to be asking the same broad question—namely, what explains the choice of certain identities as opposed to others—the scale of observation is clearly different. In *HC*, Ile Ife was studied in depth and an assumption was made that it was a microcosm of the entire Yorubaland. In *IF*, unlike *HC*, four countries were observed. Surveys inevitably had to be given greater weight than ethnography. Hopf seems to suggest that Laitin should have done in Narva, his base in Estonia, what he did in Ile Ife, but seeing all of Estonia through the prism of Narva, *let alone three other societies*, was not the purpose of Laitin’s research. Nor might it have been a sensible methodological strategy.³

This does, however, lead to an important question: can surveys be designed in such a way that they pick up some of the strengths of ethnography? A fuller discussion of this issue will lead us too far in a cognate area. It will suffice to note that making the survey questions about ambiguities, anxieties, fear and hopes—emotions that so often accompany identity poli-

tics and in quite intense forms—open-ended, collecting narratives about them and postcoding them (once narratives have been collected) is perhaps one of the best ways to go. This survey strategy is different from following a theoretically determined finite-answers form and, therefore, a precodable format, as is typical of standard surveys. I am currently experimenting with such survey designs in my own research in four countries. The upcoming results will show how far the redesigned survey technique works. Basically, those who survey do not collect narratives, and those who collect narratives do not survey, but there is no theoretical reason to see them as irreconcilable methodological adversaries. They can be substantially combined, blunting the edges of the tradeoff.⁴

Methods and Explanations

This said, another side of Hopf’s methodological critique remains. Following his point about how method is deployed in *HC* as opposed to *IF*, one could also see the basic change in Laitin’s position on what structural transformations do to human choices. In *IF*, Laitin argues that after a structural transformation brought about by the fall of the Soviet Union, the Russian-speakers in the Near Abroad calculated whether linguistic assimilation was in their interest or not. In *HC*, Laitin had said something dramatically different. The Yoruba did not calculate, when faced with the clear possibility of structural transformation in their political arena. The fascinating question for Laitin’s inquiry during his Ile Ife field work became the following: why would the Yoruba still stick to a tribal (ancestral city) identity rather than a religious (Muslim vs. Christian) one, even though a Civil War in Nigeria during 1967-70 attempted to redefine Nigeria into a Muslim North and Christian South, and again, when in the late 1970s, a debate on whether there should be a Federal Sharia Court of Appeal sought to do the same? In their religious life, the Yoruba acted as Muslims and Christians, but they remained politically committed to their tribal identity, refusing to react religiously to the cataclysmic political events. Why?

Laitin explains why rational choice is unable to help him answer this question:

Rational choice theorists...cannot tell us if ultimately butter is better than guns; it can tell us that at a certain point the production of a small number of guns will cost us a whole lot of butter, and at that point it is probably irrational to produce more guns. Within a political structure, individuals constantly make marginal decisions. [Rational choice] theories can give us a grasp on how individual political actors are likely to make choices within that structure.

[Rational choice] theory cannot, however, handle *long-term and non-marginal* decisions. When market structures are themselves threatened, and people must decide whether to work within the new structure or hold on to the old—without an opportunity for a marginal decision—microeconomic theory is not applicable...Structural transformations—changing the basic cleavage structure of a society—are not amenable to the tools of micro-eco-

conomic theory... (*HC*, 148-9, parenthesis and emphasis added).

Identity choice was not a marginal, but a structural decision. Rational choice arguments, therefore, were inapplicable.⁵ What would apply instead?

...Gramsci provides the solution....The model of hegemonic control can help explain the reification of the “tribe” in African politics—why that cleavage became the dominant metaphor for political action and why it persisted... (*HC*, 150).

To explore how hegemony was created, Laitin then goes into history, fixing his gaze on the British colonial period, starting in the late 19th century:

Claims based on religious identity were expunged from the political arena by British administrators...British administration shied away from the promotion of Christianity... British administrators...feared the revolutionary implications of religious fanaticism (*HC*, 154).

Finally, Laitin sketches the impact of this decades-long principle of British rule on the Yoruba:

The idea that ancestral city represents ‘blood’ while religion represents ‘choice’ is so deeply embedded into commonsense thinking that experience and data demonstrating otherwise fail to disabuse Yoruba people of this ‘truth’ (*HC*, 159).

This is fascinating puzzle-solving. In many parts of the world, religion is often not seen as a matter of choice, even though it is in principle. Religion is more often seen as an unchangeable reality inherited from forefathers. Moreover, in other parts of the British Empire, the colonial authorities chose religion as a ruling strategy, for example in Northern Nigeria, but in Yorubaland, they chose a different strategy, leaving a quite different legacy. The distinctiveness of the institutionalized commonsense of Yoruba politics, Laitin argues, is thus linked to the contingencies of colonial rule.

Let us now ask how the impact of a structural transformation on identities is handled in the Near Abroad. The unraveling of the Soviet system is in many ways conceptually analogous to the Biafran Civil War. It ended a system as it existed, without making it clear what would replace it instead. In *HC*, “the politicization of communal identities cannot be fully understood by examining the logic of individual choices” (103). In *IF*, whether or not Russian speakers assimilate is based on a strategic interaction with other Russian speakers, conceptualized as a matter of individual choice in a tipping model. In *HC*, calculations about identity, if made at all, were thought to take place in normal times, not in times of structural transformation, for the latter was marked by a radical uncertainty about the future, making a cost-benefit calculus hard to practice. “The level of costs and benefits of different forms of political identification among the Yoruba is not at issue. For a Yoruba to reformulate his political identity on the basis of his religion

would involve great uncertainty” (*HC*, 149). In *IF*, Russian-speakers calculate even in times of great structural transformation—namely, the collapse of the Soviet Union.

The remarkable difference between the two methods is thus clear. In *IF*, Laitin starts with an apriori theory of individual choices, as opposed to viewing choices as embedded in a structural context, as was the case in *HC*. But is it necessarily a problem? Is something lost in the process? Hopf (2006) is sure about the great loss, as discussed below. I would like to answer the question in two ways, one of which goes in the direction of Hopf’s observation, but the other does not.

Hopf argues that if Laitin had allowed himself to be an ethnographer *a la HC*, he would have found that the absence of Russian ethnonational mobilization in the Near Abroad was consistent with a variable “omitted” from the surveys: Soviet identity. Decades of history had made Soviet identity a lived everyday reality for Russians in the Near Abroad: “In understanding themselves as multinational, as Soviet, the ethnonational axis of identification was simply not available, or salient, to millions of Russians living abroad. Therefore, they did not understand themselves in opposition to Kazakhs, Ukrainians and Estonians, in precisely that way, just as the Yoruba did not mobilize along Christian and Muslim axes of identification” (Hopf 2006, 18).

Notice the role structural context plays here in the exercise of individual choices. Some choices are simply not part of the institutionalized commonsense of politics because of how history played itself out. Hopf’s central methodological insight is that *an a priori theory led Laitin to formulate his survey questions in a way that ruled out this explanatory possibility, and an ethnographic soaking and poaking would have made it transparent.*

If true, this is a very big conclusion, for it not only changes how we explain the absence of Russian ethnonational mobilization in the Near Abroad—as a result of each Russian calculating how other Russians will behave, as Laitin does, or as a result of a historically produced choice pattern, as Hopf proposes—but it also shows that an important potential substantial explanation is eliminated by a method relying on apriori assumptions.

In the end, the area experts will have to judge the veracity of either claim. What those of us doing surveys in different parts of the world can do is ask whether questions about a possible Soviet identity were included in Laitin’s questionnaire—especially in an open-ended form which allows one to watch against excessive theoretical determination of survey questions. The way Laitin’s survey questions are reported in the appendix of *IF* does not make it clear whether he did ask questions about the possibility of a Soviet identity of Russians in the Near Abroad, and in what form.

If Hopf is right, he tellingly shows us the consequences of a method driven by a priori assumptions, but I also wish to argue that the same method has also generated some big ideas. EIC by Fearon and Laitin is another example of an argument based on apriori and universalist assumptions. It proposes “in-group policing,” or “self-policing,” as a societal mechanism of peace between diverse ethnic groups. The idea is de-

ductively laid out, and Hopf is right that the theory is not empirically tested.⁶

But the fact remains that it is a big and novel idea in the field, and it has a huge empirical potential. The existing theories were either primordial (ancient hatreds), instrumental (political entrepreneurs mobilizing ethnicity for self-serving ends), epochal (arrival of modernity), or institutional (consociational or liberal democracies; voting systems, etc).⁷ Using a simple insight that ethnic groups can monitor their own group members much more easily than those of a different ethnic group, Fearon and Laitin turned it into a serious theoretical proposition, elaborated with game theory.

I am empirically testing this theoretical idea in my current project in 18 cities across Indonesia, Malaysia, Sri Lanka, Nigeria, and India, and the chances that some, if not all, of my materials will bear it out are very high. My previous arguments based on six cities in India had proposed a different societal mechanism of peace–interethnic civic engagement, especially in organizations (Varshney 2002). As my research moves further, we will perhaps find out the conditions under which in-group policing works as a mechanism, and conditions under which interethnic engagement does. In short, even though the theory that I will develop will not be universal, it is the universalist assumptions and a deductive mode of theorizing that produced the idea of in-group policing. Fearon and Laitin are certainly taking the world of knowledge forward.

How does this discussion relate to my central argument? The same method that produced in-group policing as an idea *perhaps* managed to rule out, if Hopf is right, the possibility of a Soviet identity for the Russians in the Near Abroad. And the method that identified the role of colonialism in producing institutionalized commonsense in Yoruba politics cannot easily tell us why despite British attempts at creating or freezing an institutionalized Hindu-Muslim divide through electoral rules in India, South Indians managed to escape *Hindu-Muslim* cleavages, instead getting *intra-Hindu* caste divisions as the master narrative of politics (Varshney 2002). Recourse to colonial practices resolves an intriguing puzzle about Yoruba politics and generates a possible idea—contingencies of colonial rule—for portability, without clinching it for all postcolonial analytic sites. Historians had begun to zero in on this idea elsewhere, but political scientists on the whole had not.⁸

Arguments and Statistical Testing

Let us now turn to statistical methods such as regression analysis, and examine the idea of tradeoffs. In EICW, Fearon and Laitin advance the argument about why civil wars occur in a very important way. A widely discussed recent theory, also based on statistical testing, had proposed that the odds of civil war were strikingly correlated with primary commodity exports (Collier and Hoeffler 2001). Primary commodities are “lootable” commodities, and it is “greed” about these resources that drive an insurgency, not “grievances” about ethnic discrimination, argued Collier and Hoeffler. Fearon and Laitin disprove this argument conclusively.

But they run into some trouble when they test another argument—the so-called modernist view of nationalism asso-

ciated with Anderson (1983) and Gellner (1983), both of whom attribute the emergence of nationalism to the rise of modernity, and claim that nationalism was impossible before the modern age, though their mechanisms are somewhat different (printing press and capitalism for Anderson, and industrialization for Gellner).

For a statistical testing of the modernist argument, Fearon and Laitin needed variables that could measure modernity, or proxy for it. Higher levels of per capita income became their proxy for modernity (Fearon and Laitin 2003, 78), and they find that lower levels of per capita income increased the odds of civil wars, not the other way round (83). Anderson and Gellner, they concluded, were wrong.

Were they? There are two conceptual problems in the way Fearon and Laitin formulated the test, partially inescapable due to the requirements of regression analysis. First, both Anderson and Gellner made *epochal* arguments—arguments that focused on a transformation of human consciousness as it existed in the Middle Ages, once modernity arrived. Higher or lower income of countries today—or since 1945—is quite beside the point. Pre-modern times may have had lower per capita incomes than modern times, but Anderson and Gellner also talk about print capitalism and industrialization. *Their arguments are historically specific.* The only way to test their arguments is to explore whether before the birth of the printing press and/or industry, national consciousness existed.⁹ Second, the arguments of Gellner and Anderson are about national *identities*, not nationalist civil wars. Having a national identity does not necessarily imply a hunger for war. Identities and wars are conceptually separable.

Thus, regression analysis is a good way to test some theories, not all. Contemporary primary exports are easily quantifiable at a large-n level, but how does one quantify the extent of printing press penetration in a large number of cases in the 17th, 18th, and 19th centuries, and their consequences for human consciousness? And if a large-n dataset cannot be created, how can one run regressions? In the absence of large datasets for the 17th through 19th centuries, one will have to discover a *small number of critical cases* that show the rise of national consciousness before the birth of the printing press and/or industry. The debate between Gellner and Kedourie goes precisely in that direction (Kedourie 1993, 136-144).

I hope it is now clear why we should not damn methods as intrinsically superior or inferior, neither ethnography, nor surveys, nor for that matter deductive work, whether conceptualized formally (as in game theory) or informally (as in the writings, let us say, of Rawls or an Elster). We should simply recognize the potential and limits of each method, and we should see whether the method proposed is suitable for the problem at hand.¹⁰

This is true even in the natural sciences. Einstein famously argued several decades back that physics and meteorology will neither have the same methods, nor the same degree of predictiveness. Physics typically studies a few variables in interaction, allowing parsimony and predictive accuracy. Meteorology has so many variables that having more powerful computers, which he saw coming, would only allow us to

predict whether a broad area would get hurricanes in the week or so ahead, not predict that months in advance and if closer to time, not predict whether a specific town or village would be hit and with what intensity. Einstein's reasoning was clear: it is the number of variables affecting the path and intensity of a hurricane (or a snow storm) and their very complicated interaction that was at issue here, not our computing powers. Meteorological problems cannot be reduced to a few variables, as in physics. Likewise, some problems in politics may well allow the parsimony of physics, but others may be more like meteorology, requiring very different kinds of conceptualization and measurement.

I should add that my argument about methodological choices entailing trade-offs is consistent with some new work on methods, both of the quantitative and qualitative sort (Brady and Collier 2004). It also underlines the value of methodological pluralism in the social sciences. Methodological pluralism is defensible not because anything goes, but because different methods will do different things well.¹¹

Further Implications

For Laitin, this argument has some further implications. His movement from ethnography and case studies to surveys, formal reasoning, and statistical testing should allow him to deal with puzzles of a large variety. At the same time, his denunciation of other people's work in his more recent scholarly phase is puzzling and hugely paradoxical. Consider two examples, one about case studies, another about selection on the dependent variable.

Laitin finds case studies unacceptable unless the study of a single country is "transformed into a high-n research design, thereby increasing the scientific leverage" (Laitin 2003, 180). My argument with this reasoning is not that turning a country study into a high-n design is wrong. Rather, I have problems with Laitin's insistence that that is the *only way* to save case studies, or by extension, ethnographies, which tend to study a village or a town.

Paradoxically, Laitin's argument today amounts to denouncing his old scholarly self, so evocatively in evidence in *HC*. More generally, Laitin's insistence ignores the fact that case studies can contribute to cumulation by producing intriguing ideas, even when the n is equal to one. This is true of *critical cases*, which even the more statistical view of King, Keohane, and Verba (1994) accepts as valid. To recall, critical cases are those that, given theory, you would least expect to have outcomes that they do. With extensive low incomes and widespread illiteracy, India should not have been democratic, but it is. With little sanctity of private contracts, virtually no restrictions on the powers of the state, and a highly underdeveloped capital market, China should not have been an economic dynamo for two and a half decades, but it has been. India and China thus become critical cases for the theories of democracy and economic growth. Theoretically unselfconscious case studies are a problem, not country studies that are not transformed into a high-n design.

Laitin's critique of studies that select on the dependent variable is also oddly self-defeating (Laitin 2003, 179). It ig-

nores, first of all, the value of his own *IF*, where he studies identity formation only in those parts of the former Soviet Union where conflict was absent or low in the immediate post-Soviet phase: Kazakhstan, Estonia, Latvia, and Ukraine. He does not study the Chechen region, which had a lot of conflict and could have had very different identity outcomes.

Moreover, some of the most instructive social science work in recent decades selects even more on the dependent variable than Laitin does in *IF*. Sen's *Poverty and Famines* (1981) and Bates' *Markets and States in Tropical Africa* (1981) are the best examples. Both are widely viewed as classics of the development field, and justly so. Sen's theory of famines was based on five famines; there were no half-famines or non-famines in his research design. And as Rogowski pointed out long ago (1995), Bates only studied agricultural stagnation in Sub-Saharan Africa, not cases of agricultural success.¹²

It may be true that in most cases, it will be hard to clinch an argument we want to make if we have no variation on the dependent variable. But that is not the only way to contribute to knowledge. Both Sen and Bates did two notable things. First, they thoroughly undermined an existing conventional wisdom: food availability decline as a cause of famine, and Africa's cultural taste for leisure over work, producing backward bending supply curves instead of upward sloping ones, leading to agricultural stagnation.¹³ Second, they put a new idea on the table for others to work with: entitlement failures as a cause of famine, as in Sen's argument, and self-seeking behavior of urban politicians, buying the rural rich through subsidies, and running policies that hurt the countryside as a whole, as in Bates.¹⁴

Research designs that select on the dependent variable can often do both of these, and that is reason enough to see them as contributions. Clinching theories in an ideal fashion is one way to contribute; undermining existing popular theories and presenting elements of a new are another way. Interestingly, in *IF*, there is a point where Laitin says something similar (Laitin 1998, 325), but he nonetheless attacks such studies elsewhere for they do not contribute to cumulation (Laitin 2003, 179).

Conclusion

Laitin has made remarkable contributions to our knowledge, becoming a central figure in the subfield of ethnic politics. The kind of methodological evolution he has undergone is also uncommon in the profession. For both of these reasons, substantive and methodological, his scholarly output inspires admiration. The admiration would be infinitely greater if he could view methodological choices as consisting of tradeoffs and could thereby view work emanating from methods not currently favored by him as also contributing to the cumulation of knowledge, for which the intellectual case is quite clear. It will also save him from self-inflicted paradoxes and contradictions.

Notes

¹ I would like to thank Anna Grzymala-Busse, Ira Katznelson, David Laitin and Daniel Posner for some penetrating comments on an earlier version of the argument presented here.

² For a fuller development of this distinction, see Varshney (2003b). See also Isaiah Berlin, 1979.

³ Whether Laitin should have engaged in four ethnographies, selecting a central site in each country, is an important issue, and worth thinking about. But studying all four societies from the microcosm of Narva would not have been methodologically valid.

⁴ For some early thoughts on these lines, see Varshney, 2002, 19-20.

⁵ Whether this critique of rational choice theory is right is a different point altogether. On what kind of rationality might apply to non-marginal decisions like identity choices, see Sen (1982) and Varshney (2003a).

⁶ Only examples of an approving sort are listed by Fearon and Laitin. See the empirical critique of Horowitz (2001, 475-6).

⁷ A detailed elaboration of these traditions can be found in Varshney (2002), Chapter 2.

⁸ With the exception of Benedict Anderson (1983), especially in his account of "Creole Pioneers."

⁹ Alternatively, can better proxies be developed for epochal arguments? One should, of course, remain open to such possibilities.

¹⁰ In one of his recent essays, Laitin (2003) appears to have partly moved in this direction. He argues for a tripartite method: formal reasoning, statistical testing, and narratives. But it is unclear whether the ideal set forth has ever been realized, or can be. Moreover, considerable paradoxes in that position also remain, as discussed later.

¹¹ On this matter, also see Laitin (2003) for a different view.

¹² To be fair to Bates, he does mention Kenya and Ivory Coast as cases of relative agricultural success, but that account comes at the end and is very brief. Basically, variation in outcomes is not the centerpiece of the argument.

¹³ See Varshney, 1995, Ch. 2, for how popular the theories of backward bending supply curves were.

¹⁴ This idea did have a prior lineage in Lipton (1977) and Schultz (1980), but Bates provided the most convincing links between politics and economic outcomes. Lipton and Schultz assumed that the urban bias of the political structure produced anti-agricultural outcomes. Bates showed exactly the links worked.

References

- Althusser, Louis. 1969. *For Marx*. London: Allen Lane.
- Anderson, Benedict. 1983. *Imagined Communities*. London: Verso.
- Bates, Robert. 1981. *Markets and States in Tropical Africa*. Berkeley: University of California Press.
- Berlin, Isaiah. 1979. *Russian Thinkers*. Henry Hardy, ed. New York: Penguin.
- Collier, Paul and Anke Hoeffler. 2001. "Greed and Grievance in Civil War." World Bank, Typescript. <https://econ.worldbank.org/programs/library>.
- Dahl, Robert. 1982. *Dilemmas of a Pluralist Democracy*. New Haven: Yale University Press.
- Elster, Jan. 2000. "Rational Choice History: A Case of Excessive Ambition." *American Political Science Review* 94:3 (September), 685-95.
- Fearon, James and David Laitin. 1995. "Explaining Interethnic Cooperation." *American Political Science Review* 90:4 (December), 715-35.
- Fearon, James and David Laitin. 2003. "Ethnicity, Insurgency and Civil War." *American Political Science Review* 97:1 (February), 75-90.
- Gellner, Ernest. 1983. *Nations and Nationalism*. Ithaca: Cornell University Press.
- Hopf, Ted. 2006. "Ethnography and Rational Choice in David Laitin: From Equality to Subordination to Absence." *Qualitative Meth-*

- ods: Newsletter of the American Political Science Organization Organized Section on Qualitative Methods* 4:1 (Spring), 17-20.
- Horowitz, Donald. 2001. *The Deadly Ethnic Riot*. Berkeley and Los Angeles: University of California Press.
- Kedourie, Elie. 1993 [1961]. *Nationalism*. Oxford: Blackwell Publishers.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton: Princeton University Press.
- Laitin, David. 1977. *Politics, Language and Thought: The Somali Experience*. Chicago: University of Chicago Press.
- Laitin, David. 1986. *Hegemony and Culture: Politics and Change Among the Yoruba*. Chicago: University of Chicago Press.
- Laitin, David, 1998. *Identity in Formation: The Russian Speaking Population in the Near Abroad*. Ithaca: Cornell University Press.
- Laitin, David. 2003. "The Perestroika Challenge to Social Science." *Politics and Society* 31:1 (March), 163-184.
- Lipton, Michael. 1977. *Why Poor People Stay Poor: Urban Bias in World Development*. Cambridge: Harvard University Press.
- Rogowski, Ronald. 1995. "The Role of Theory and Anomaly in Social Scientific Inference." *American Political Science Review* 89:2 (June), 467-70.
- Schultz, Theodore. 1980. *Distortion of Agricultural Incentives*. Bloomington: Indiana University Press.
- Sen, Amartya. 1983. *Poverty and Famines*. Oxford: Clarendon Press.
- Sen, Amartya. 1982. "Rational Fools." In *Choice, Welfare and Measurement* (Cambridge: MIT Press).
- Varshney, Ashutosh. 1995. *Democracy, Development and the Countryside*. New York: Cambridge University Press.
- Varshney, Ashutosh. 2002. *Ethnic Conflict and Civic Life: Hindus and Muslims in India*. New Haven: Yale University Press.
- Varshney, Ashutosh. 2003a. "Nationalism, Ethnic Conflict and Rationality." *Perspectives on Politics* 1:1 (March), 85-99.
- Varshney, Ashutosh. 2003b. "Varshney and Bates; Two Views on Seeing Like a State." *APSA-CP* (Summer), 1492.

Ethnography and/or Rational Choice: A Response from David Laitin

David Laitin
Stanford University
dlaitin@stanford.edu

As Ted Hopf presented his paper at the symposium held at the 2005 annual APSA meeting, provocatively titled "Being David Laitin," I felt as if I were in a chute on the 7^{1/2}th floor of the Marriott Wardman Park, ready to be discharged onto the New Jersey turnpike. But I survived, enough so to offer the following remarks.

The key substantive theme raised by the papers in the symposium is the relationship of ethnography and a theory of purposive action. In the 1950s, the eminent anthropologist Frederic Barth encountered the work of John von Neumann and Oskar Morgenstern, and immediately saw the deep implications of their game theory for anthropology. He then wrote a game theoretic essay (Barth 1959) analyzing chieftaincy politics among the Pathans. This was one of the few lead balloons that Barth let fly in his distinguished career, and the anthropological field has steered clear of game theory ever since. But the scholarly relationship between game theory and ethnogra-

phy remains a promising one, and I was delighted that Ted Hopf identified this complementarity as an opportunity for *our* discipline to grasp. Indeed, many issues need to be worked out if this project is to become part of the *habitus* of comparativist research; but this symposium has made a significant step in exploiting this opportunity.

A subsidiary theme in the papers explores the relationship of ethnography and statistics. Since I have advocated for in comparative politics what I call the “tripartite method”—one that includes ethnography (one exemplar of narrative), formal models, and high-n research, the symposium gives me the opportunity to address the criticisms of this research strategy that were raised in several of the papers.

In the first section of this response, I will address substantive themes. The panel papers offer challenges to the methods that I have practiced and advocated. Yet I remain confident that research in comparative politics that relies on the tripartite method—more or less as I have outlined—remains both general and attractive. In the second section of this response, I will address some specific criticisms of my research. I should say in advance, despite the critical stance that I shall adopt, that I take the seriousness of the criticisms in this panel as high praise. And to get credentials as both a hedgehog *and* a fox, awarded by Varshney, is more meaningful to me than an honorary degree!

I. The Substantive Issues Raised in the Panel: Ethnography vs. Game Theory

The notion that there is a trade-off between ethnography and theories of purposive action (of which game theory is a member) obscures the opportunity for complementarity. Consider the following formulation. Varshney writes that Laitin’s “methodological commitments have radically changed over time” and describes “Laitin’s methodological voyage from his early work based primarily on ethnography, whether in Somalia or in . . . Nigeria, to his work over the last decade and a half, in which a rational choice stance has played a major role.”

This formulation assumes that ethnography and rational choice are antithetical methods, such that in any research project one needs to subscribe to one or another of these two methods. But this misses their complementarity. Ethnography is a method of collecting and reporting narratively on social and cultural data in circumscribed communities. It is a member of the set of data gathering activities that includes archival research, surveys, and interviews and permits the reconstruction of events and processes. Rational choice is a method of analyzing purposive behavior. It is a member of the set of research activities that includes game theory, simulation, and prospect theory. It permits an analytic focus on key causal mechanisms that drive social outcomes. There is no choice between two methods here; rather, ethnography and rational choice play complementary roles in the logic of discovery.¹

In my research, the source of insight has almost always been ethnographic. What has changed is the level of analysis in which I have sought explanation. In my early work on Somalia and Yorubaland, a macro-sociological focus led me to examine factors such as bureaucratic (in maintaining European lan-

guages) and colonial (in structuring incentives for group expression) interests. In my more recent work on Catalonia and the former Soviet Union, a microanalytic focus led me to examine factors such as individual incentives to assimilate and to rebel.

In comparing my denial of rational identity calculation in Yorubaland and my emphasis on it in Narva, Varshney writes, “the remarkable difference between the two methods [ethnography and rational choice] is thus clear.” I agree with Varshney that there was a change in my orientation across these two projects, but it was not one of embracing (in the Yoruba case) and abjuring (in the Narva case) the ethnographic method. The macro focus on social forces and the micro focus on individual action were both constructed on an ethnographic base. In sum, in all of my four extended field exercises—Somalia, Yorubaland, Catalonia, and Estonia—ethnography informed a variety of theoretical orientations.

Varshney, furthermore, sees the micro/macro trade-off in too-stark terms. In a post-dissertation paper on Somalia (Laitin 1983), I reexamined my formerly macro analysis of Somali nationalism from a micro perspective, foreshadowing the change in focus that marked my early papers on migrants to assimilate in post-Franco Catalonia. In my subsequent work on Yorubaland, in exploring the “two faces of culture,” I analyzed the microincentives of Yorubas to use ancestral city attachments, however constructed, for political gain. And in *IF*, where I focused mainly on microdynamics, I did not ignore the macro incentives of Soviet structures. In fact, I compared the colonial, the most-favored-lord, and the integralist models of political incorporation in chapter 3. These three historical experiences of political incorporation set the stage for the differential pay-offs in the assimilationist tipping game that the Russian-speakers “played” after the Soviet collapse.

The complementarity of ethnography and game theory is demonstrated by the degree of dependence game theory has on truly thick ethnographic description. Consider Woodruff’s re-analysis of what he calls the *locus classicus* of ethnography, Geertz’s “Thick Description.” On the two incidents in Morocco involving the Berber traders as described in that essay, Woodruff offers a counter-interpretation, consistent with Geertz’s data but quite different from Geertz’s (and Fearon and Laitin’s 1996) interpretation of the events. As opposed to Geertz’s focus on the institutional incentives afforded by the trade pact, Woodruff points out that “when Geertz’s sheikh had seven relatives killed as punishment for a trivial theft, the formal rules of the trade pact may well have provided a convenient occasion or pretext for a display of dominance needed for other purposes.” Woodruff points out that the data do not allow us to choose because “there is nowhere near enough evidence to sustain either of these interpretations—but there’s not enough to impugn them, either.” Woodruff is right! The details Geertz offers in his self-proclaimed “thick description” are insufficient to construct a unique game model that captures the nub of the strategic interactions. I infer from this that game theory is hungrier for details than the community of scholars that granted Geertz’s essay with the status of a descriptive *locus classicus*.

Thinking about the hunger for detail that any formal model requires, I concede to Varshney that I insufficiently explored the microfoundations of identity stability in the Yoruba case. Note well that I did not ignore them. I explored the several attempts by identity entrepreneurs to move from an ancestral city to a religious equilibrium, and reported on them as failures. But I took those failures too much for granted. In retrospect I should have pursued the matter more forcefully, trying to figure out why Yoruba Muslims from weak ancestral cities (i.e., those ancestral cities that conferred little status) did not attempt to foster an identity cascade towards a religious orientation (where their status might have been higher). Working only in Ile-Ife, where ancestral city status was high, I ignored the incentives of aspiring politicians in low status cities. Such an exercise would have demanded closer attention to what game theorists call off-the-path expectations (such as what Yoruba Muslims think will happen to their social status and their daughters' marital worth if they join in an alliance with Hausas). Ethnography written without a formal model (one that the Gramscian perspective that I adopted lacked) is likely to ignore such off-the-path expectations that drive everyday behavior. The formal model (Schelling's tipping game) that I used in *IF* pushed me to ask interesting questions on the micro level that were not fully explored in *HC*. A microanalytic game theoretic model would have pushed me to do better, deeper, ethnography in Yorubaland than what I actually did.

To be sure, not all game theoretic models are "thick" redescrptions of ethnographic reality. There is a class of models that are thin, that seek to capture decontextualized generic processes.² One such model gets Hopf's goat. He writes that in "Explaining Interethnic Cooperation" (EIC) written with James Fearon, "intersubjective reality disappears as an object of theoretical or empirical interest. All social phenomena are either objectively labeled, or assumed to have a particular value. This article received an APSA prize as the best article to appear in the 1996 *APSR*. Clearly, objectivism has its rewards."³ I'm not sure what objectivism is, but Hopf is correct that EIC was a departure from *HC* in that Fearon and I decontextualized a social mechanism in order to isolate its universalist implications.

Hopf goes on to write: "But why assume the need for iteration...Isn't the assumption of a need for iteration here a move dictated for model specification, rather than one grounded in consensual empirical research? And isn't it precisely the opposite of what Laitin did in *HC*?"

The assumption that iteration is a mechanism affording intra-ethnic trust relations was indeed a modeling decision, but one that is defensible. The then-reigning literature on ethnicity and violence explains intra-ethnic mistrust, but cannot simultaneously explain extraordinary low levels of inter-ethnic violence. Our model, by interpreting *intra*-ethnic relations as peaceful due to high information, provided a new account for variation in *inter*-ethnic violence. As it stood in 1996 it was only a conjecture, albeit given some plausibility through a look at ethnographic accounts of inter-ethnic relations. It would require an empirical project—of the kind Varshney describes in this symposium, but also in the systematic tests of

its observable implications—before it could be taken as anything more than a formalized conjecture. In this case thickness would follow the abstract model rather than inform it. But my general point is that connecting a model to thickly described choice frameworks is a powerful tool; in no way is a thin model a substitute for the details of social and political relations.

Even if ethnography feeds game theory, perhaps, as suggested by Herrera, there is only a limited class of politics for which game theory can effectively complement ethnography. Herrera writes: "And most importantly, there are classes of problems for which formalization is not appropriate (yet, or perhaps ever): these include issues such as irrational beliefs and behavior, non-transitive preferences, interactions in which new unknown and unknowable possibilities for action exist. Formal theory is not able to solve these sorts of problems at the moment, and may never be able to."

This is unconvincing. Take suicide missions, an arena of activity that is fraught with irrational beliefs. Thin game theory has been particularly reprehensible in this area. We learn nothing by adding a slew of virgins in heaven onto the attackers' utility functions to make the activity appear rational. But to say that the actions are based on "irrational beliefs" or "non-transitive preferences" is to give up the game of analysis too quickly—and it too often leads to tautology. If we ask why Hamas is good at suicide attacks but not Fatah; or why Palestinians can get recruits for this but not Irish Catholics; we are compelled to model incentives, beliefs about others' beliefs, and alternative opportunities for rebels. Recruits for suicide missions may be driven to enlist in order to bestow status on their families; but then find at the moment of truth that defection is impossible. A good model would show *inter alia* the conditions under which defection is less likely—and under those conditions we would expect to observe the irrational fulfillment of what might have initially appeared as a non-credible commitment.⁴

Along these lines, Herrera and Woodruff both insist on limits to game theory to cases where every player is motivated by the same set of rules. Woodruff writes "even those who accept rational choice as a working assumption should conclude that game-theoretic methods of *describing situations* have little advantage when there are distinct contextual incentives for conformance to rules in distinct circumstances." But I would argue that if the nub of a social situation entails players conditioning their behavior on different signals, this is what should be modeled. Suppose I shoplift only when I know that surveillance is weak; but Woodruff never shoplifts because he has strong honesty norms. In Wal-Mart there would be a "pooled equilibrium," neither of us shoplifting, because surveillance is strong. But a model that links our beliefs to our strategies would predict that if Wal-Mart's internal police were to go on strike, Laitin but not Woodruff would garner the spoils. A good model would show that under surveillance conditions, Woodruff and I would be honest but for distinct reasons; the model would be confirmed with a behavioral divergence if there were a change in our expectations about getting caught. The model would be showing distinct

contextual incentives for conformance to rules.

To concretize his point, Woodruff offers an example of contracts and weaseling out of them. The context of the contract is so dense that any number of possibilities can emerge. In fact, he writes, "Betty's bill-paying behavior will reflect *multiple* equilibria in *multiple* games, in which the rules of contract are simply a single element. Some of these games will differ only in the payoffs to different actions; others will feature moves, such as shifting suppliers, not always available." Woodruff expects game theory to make a point prediction about a particular contract when there are more than several parameters. This is an extraordinary demand. It would be like asking a physicist to predict the exact moment a feather dropped from the Eiffel Tower will hit Parisian turf. If Woodruff were to ask (sticking to his example) the variety of moves available to a future's buyer when the price of the good she contracted for drops radically in value, game theory might offer some help in organizing the parameter conditions under which weaseling would be the optimal response. Then we could collect a dataset of different future's markets to see if the model worked as theorized. If it did not, further theorizing would be called for; if it did, we would have confirmation of a rudimentary theory of weaseling behavior in future's markets.

Woodruff insists, however, "In the...sorts of cases [where rules may be formally applied, or informally ignored], game theory, I think we'd have to admit, isn't very powerful. All it amounts to is a kind of a protocol for writing down the results of ethnographic research in what could be a huge number of different models." Woodruff here misses a big point, viz., that game theory calls out for thick ethnographic information as to when a rule will be applied and when it will be subtly ignored. Modeling that choice point would be an exciting opportunity for our general understanding of the boundary separating formal from informal institutions.

Game Theory vs. Statistical Tests

Statistical tests of relationships uncovered in ethnographic settings were also challenged at the symposium. One challenge is if such tests could ever set the proper controls embedded in case studies without losing all degrees of freedom. Another challenge demands that statistical tests, to be useful, would have to be conclusive.

In making the claim that statistical tests would never capture local controls, Chandra offers a fascinating brief for mechanisms, the analysis of which she claims can make only within-case point predictions. She introduces an innovative method of mechanism-tracing to reinterpret my work on the emergence of politicized cleavages. She sees the fruits of this approach as fulfilling the ultimate goal of social scientific explanation, one that does not require statistical tests of each link on the mechanistic chain of micro-causes. Indeed, she holds little hope for statistical tests. She asks: "Could Laitin have tested the 'correct' model [of colonial imposition on cleavage structure]? Could anyone, on the basis of a within-country study?" She thinks not, as the problem of embedding controls would yield insurmountable degree of freedom problems.

Chandra's pessimism is not fully justified. In the case at

hand, consider a cross-sectional model with precolonial social cleavages, the colonial strategy in regard to cleavages, and in their interaction as independent variables. The post-colonial dominant political cleavage would be the dependent variable. If colonial strategy were not significant, the result would point to a neat anomaly from the theory offered in *HC* and that further research would be needed to answer. If, however, there were a significant marginal impact of colonial strategy on post-colonial politicized cleavages, then the statistical model would have added some provisional general knowledge beyond the case study. The theory will never be proven—as Popper has shown, that never happens in positivist research—but a statistical test with significant results would add confidence that the colonial hegemonic model correctly described the sources of political cleavages in the postcolonial world.

Chandra takes an even more radical stand against cross-sectional statistical tests of within-country findings. She writes: "Indeed, while we can think of numerous examples of within-country studies that identify interesting, internally consistent, bivariate hypotheses, it is difficult to think of bivariate hypotheses generated from within-country studies that have actually been verified through cross-national research."

This claim can be challenged theoretically and empirically. Theoretically, even if case studies hide variables through assumptions about context that cannot all be included in cross-national statistical tests (as Chandra argues), we learn something if the variables that appear significant in our case studies work as well in cross-national tests. If significant in high-n tests, we learn that the variables identified in the case study are robust across contexts, adding to our confidence that the identified variable is broadly significant.

Empirically, Chandra's claim is provocative, and will require responses from the wider research community. But I might note as a counter-point that a generation of within-country research (in which Chandra has been a leading practitioner) showing the endogeneity of ethnic identities to political opportunities has been verified in cross-national research that now shows controlling for GDP, ethno-linguistic fractionalization cannot account for failures in democracy or in maintaining social order. If proof is the goal, Chandra will win this debate. But I have lower expectations. Statistical tests in this case added to confidence in our case-based findings.

In another tack to discredit large-n tests of ethnographic reality, Hopf complains that the specifications of variables necessary for statistical tests violate the richness of ethnographic insights. "Daily practices in *IF*," he argues, "are reduced to language choice. Despite repeatedly, and correctly from an ethnographic point of view, stating that language is insufficient to measure identity, to capture the broad array of daily social practices that constitute identity, Laitin ends up relying on what he has already ascribed as inadequate: language. And he does so because of his sacrifice of ethnography to surveys." Hopf is certainly correct that I took language as a proxy for culture in *IF*. But note I did the same thing for religion in *HC*, a book he upholds as a model for comparative research. The use of proxies not only makes sta-

tistical tests possible, but it allows us to get some traction on slippery concepts such as culture, even in the absence of statistical manipulation. If our goal is to explain all aspects of culture within the context of a single study, any method will drive us to indeterminate despair.

A final charge on the use of statistics to confirm ethnographic descriptions or formal models is that we have developed a cult in political science in which bad quantitative data dominate good ethnographic data. On this point, Herrera writes: "Indeed, there are those so committed to QA [quantitative analysis] that if the datasets do not exist for a given problem, they do not study it; and there are some who will resort to using any data, no matter how poor, as long as they exist." But this is a one-sided slap. Africanist historians, in the absence of archives, have long relied on oral histories which are of quite poor quality. But scholars such as Jan Vansina (1985) set standards for the use of bad historical data just as students of QA set standards to correct for a variety of biases inherent in the collection of bad data. All of us, no matter how we collect data, must be sensitive to bias; but there is nothing inherent in QA that makes it more subject to bias.

The Quest for Certainty

Common to the charges of my use of game theory and statistical tests is what I believe to be a false hope that science yields certainty. All that can be hoped for in any research activity, as I noted earlier, is to add or subtract confidence in a theory or intuition that drove us into the research. Consider Chandra's criticism of the test of my hypothesis developed from Yorubaland in the case of Benin. She writes "The book [HC] finds that the pattern of politicization in post-colonial politics can be explained by the pattern of politicization adopted by colonial rule. Consequently, Laitin argues that Benin 'demonstrates the power...of the model of hegemony.'" However, Chandra cautions, "Unless [his] entire model is exported [to Benin], we do not know whether it is corroborated or not. The fact that Benin corroborates the model, thus, is a false positive, since we do not know whether the control variables, on which the effect of colonial hegemony was contingent in Yorubaland, also took on the same value in the case of Benin."

Chandra is here looking for conclusive tests—something the natural world is quite stingy in providing—rather than opportunities to add or subtract confidence in our theories. In the case of Benin, if ancestral city had been a powerful conditional factor of political behavior among Beninois Yorubas, but influenced by a different colonial strategy, I would have lost some confidence in a model that focused on the colonial power structuring cultural cleavages. Benin wasn't a true unbiased test; rather, it was an opportunity to update my confidence in the importance of colonial strategy for post-colonial cleavage structures.

Woodruff (and Hopf as well) articulate a similar gripe in reaction to the in-group policing mechanism elucidated in EIC—namely that Fearon and I could not possibly explain most or all cases of ethnic peace with this single model. This critique confuses the full explanation for a social outcome (in which high *r*-squares are of value) and the identification of a mecha-

nism that has a consistent (if only a limited) impact on a social interaction. The Fearon/Laitin model identified a mechanism that had a marginal impact on inter-ethnic violence, but we hardly claimed that a regression model with only in-group policing mechanisms as independent variables would explain a large amount of the variance across societies.

More generally on this point, Woodruff concludes his essay by writing "it would be a mistake to build comparative politics around the assumption that tractable game-theoretic models can provide satisfactory answers to our enduring questions." I don't know anyone (certainly not I) who would make this claim. Game theory can capture the nub of a strategic interaction that models the behavior of the "switchmen [who] determine...the tracks along which action has been pushed by the dynamic of interest" (Weber 1946/58, 280)—and through the formalization of games, we can use comparative statics to design systematic tests of the model's observable implications. But it would be absurd to claim that isolating a strand of social relations is equivalent to unraveling an entire skein. Capturing the marginal effect of an *x* on *y*, and understanding its magnitude and its generality, represents good progress that we should not underestimate. It is no criticism of a research design that it will never explain everything.

The Tripartite Method

I have defended the complementarity of ethnography and game theory. I have then defended the role of statistical modeling to test the observable implications of thick game models. Thus the foundation for the tripartite method that I have advocated remains solid.

Herrera offers an amendment to this tripartite design such that the preferred menu of research options is quadripartite. Her idea to separate formal/nonformal theory, and quantitative/narrative use of data is a good one. In my paper on "The Political Science Discipline" (Laitin 2004), I advocate embedding the political theory canon into our discipline more centrally than it now stands. Herrera's amendment is consistent with my vision for the wider discipline. It would also allow me to place much of macro theory (Marxism, State Theory), formally unaccounted for in my rendition of the tripartite method, into the nonformal theory category. Yet I offer a caution. If we were to adopt Herrera's four-part division of approaches, we need to be careful to reject the apparent (or implied) affinity of nonformal theory and narrative analysis of data. One of the most promising areas of research in social science is the combination of formal theory and narrative in the sense described in *Analytic Narratives* (Bates et al. 1998) and more comprehensively in Greif (2006). Narrative at its best is in the exposition of a structure (in which the extensive form game is but one type of structure) relying on names of real actors and countries.⁵ Being able to tell a coherent narrative based on the path of play and of off-path expectations brings a new excitement to the study of politics.

Chandra goes further than Herrera in opening up a Pandora's box of approaches available to comparativists. She writes: "Ethnography and rational choice are not an exhaustive set of approaches which can be used to identify mecha-

nisms. Depending on what kinds of questions the researcher is interested in, almost any technique can be used. Those that are being used in political science so far include survey research, experimental research in the laboratory or in the field, agent-based modeling, among others—and there are many other approaches that may be imported from other disciplines, including neurobiology, psychology, psychoanalysis, even literary analysis! Why restrict ourselves to only these two?”

But the tripartite method that I advocate is not as narrow as Chandra avers. I have sought complementarities from three ways of working, each with a menu of possibilities: from formalism, there is game theory, agent-based algorithms, and prospect theory; from statistics, there are experiments, surveys, and time series cross-national regressions; and from narrative, there is history, ethnography, and intensive interviewing. In taking advantage of the complementarities across these approaches, the tripartite method pushes researchers to a deeper and broader knowledge of their subject.⁶

Allow me to return to Geertz's "Thick Description." As Woodruff points out, the Fearon/Laitin interpretation of Geertz's Atlas adventure story is no more consistent than Woodruff's, who did his without a formal model. Two points follow. First, the thicker the description, the more there are restrictions on the range of possible interpretations. A game model is thereby better founded the thicker the descriptive material that goes into its formulation. Second, formalization of one's interpretation allows for a stronger coupling between the interpretation and the observable implications of the interpretation. To the extent that the observable implications can be put to comparative statics tests, our confidence in the model grows. We will never know for sure if the stipulated motivations of the actors in the model were correctly identified; but through thick description, formal modeling, and statistical tests—that is, through the tripartite method—we can over time increase (or decrease) our confidence in our interpretation of what motivates social and political action.

II. The Criticisms of My Work and Research Program

Substantive issues aside, the panel was also about me! On that topic I have little systematic knowledge, but I can provide an historical correction, and then a defense of some of the specific claims made against my writings.

An Historical Correction

There was a premise in the organization of the panel about my intellectual development. The phrase "the Fearon factor" and the suggestion that James Fearon somehow kidnapped Laitin's brain gave pizzazz to the panel. Chairman Hopf, he later admitted, offered these inuendos in an ironic tone. But the truth is more interesting than the fiction, so I digress with a moment of autobiography. In the mid-1980s, when *HC* was in press (and, by the way, excoriated by anthropology reviewers for the University of Chicago Press and the University of California Press), I was doing field work in Catalonia, studying the breakdown of a hegemonic Castilian language regime. But fieldwork observations disabused me of this theme.

For example, the Catalan political class and many social

leaders were pressing for what they called the "normalization" of language—changing the language of everyday interaction from Castilian to Catalan. One felt in the air that *all* Catalans were waiting for generations for this to occur. The newspaper *Avui*, written in Catalan, was one example of the outpouring of nationalist fervor for a vibrant Catalan culture. It became anti-nationalist to read *La Vanguardia* (the leading Barcelona paper written in Castilian) in public, or even worse *El País*, the paper from Madrid. One day on the metro I saw a man apparently reading *Avui*, but observed from the inside, he was really reading the much more intellectual and comprehensive *El País*. I had an epiphany. Maybe normalization was not part of a wave of nationalist euphoria, but a pain in the neck that people had to adjust their lives to, given the scorn they would face should they be seen speaking or reading in Castilian. This ethnographic epiphany led me to the game theoretically illustrated mechanism that Kanchan Chandra mentions: the private subversion of public goods. (Will Rogers, an American humorist, identified this mechanism clearly: people from Oklahoma, he wrote in the 1920s, will continue to vote for prohibition as long as they can stagger to the polls.)

With this idea, I thought of the coordination games I had read under J. Roland Pennock in a political theory seminar as an undergraduate at Swarthmore College. (At UC Berkeley, where I got my Ph.D., I never heard the term "game theory." For Sheldon Wolin's Political Theory seminar I constructed a game model to compare Hobbes and Locke, but Wolin made clear that I was not speaking the right language either for Berkeley or for political theory; I stopped.) Back to Catalanization, I wracked my brain for the technology to analyze language issues from a game theoretic perspective. I then sent my ill-formed ideas from Barcelona to Robert Bates, one of the few comparativists who might have been both useful for guidance and sympathy. Two weeks later I received a response. Bob wrote that I was a good ethnographer, and it would be best for all if I stuck to ethnography. Bob is not usually considered a fan of ethnography, but he has relied heavily on the Manchester anthropologists, the work of Elizabeth Colson, and the respect and knowledge of ethnography for a rational theory of action is reflected in much of his work (but few of his public pronouncements). So Bob thought I should invest in my strengths in a field he respected. But I did not take that advice. All this to say, if Hopf had greater confidence in ethnography, he would have abstained from any suggestion that Fearon had stolen my brain.

Some Misinterpretations

Hopf has criticized (and Varshney has echoed that criticism) *IF* for ignoring the "intersubjective world of post-Soviet subjects." He charges me with "an over-reliance on a priori theory" and thereby ignoring the "set of lived daily practices" that constitutes a Soviet identity. This charge egregiously misrepresents the research process that led me to my conclusions in that book. A central finding in *IF* is that in the post-Soviet world, a new identity category emerged, that of the Russian-speaking population in the near abroad. This *idea* of the Russian-speaking population was not in my surveys;

nor was it in the formal model that I borrowed from my Catalonia research for this post-Soviet project. It was discovered by listening to the discourse of nationalism *in the field*.

In the book, I carefully analyze the role of this new identity category. I first discuss my informants' psychological connection to the Soviet identity project and how it is retained in everyday practice (91-3). Later (194), using material recorded from Latvian TV, I provide an analysis of why the Soviet identity has nevertheless "lost its luster" and is therefore absent from everyday popular rhetoric. Then (on 197-8), in the conclusion to Part II (called "An Ethnography of the Double Cataclysm"), I summarize my ethnographic evidence on the rise and fall of the Soviet identity project. It is from these ethnographic encounters that I came upon its substitute, with different implications: the Russian-speaking population. It is ludicrous for Hopf to claim that this idea came from the surveys; its magnitude and breadth was put to a large-n test (in a content analytic exercise) only because it emerged from ethnographic encounters.

Hopf misinterprets other parts of *IF* as well. "Unlike in *HC*," he writes, "where a rational choice model was informed by ethnographic evidence, Laitin allows the model here to drive the search and interpretation of evidence. Russians abroad decide to assimilate, or not, based on the relative costs of alternative strategies...But Russians' understandings of the costs are not ethnographically recovered, but instead assumed, or derived from survey questions whose formulations are already derivative of the theory being tested, not generated from the intersubjective understandings of the Russians themselves. For example, one might hypothesize that costs vary according to Russian identification with the titular nationality, such that already understanding oneself as a Russian-Estonian reduces the a priori costs of assimilation..."

In *IF*, I did not a priori theorize on the costs of assimilation. Instead, I applied matched-guise experiments that had been used by socio-linguists and anthropologists in other settings. These tests revealed a link, for example, between the historical cultural experiences in Latvia vs. Kazakhstan, and the implications for status pay-offs for assimilation. This wasn't assumed; it emerged from historical study, ethnographic observation, and then through linguistic experiments.

A second misinterpretation of my work in the papers from this symposium involves Varshney's complaints that I come down in my methodological papers too harshly on case studies, and on studies that select on the dependent variable. Varshney does not accurately portray my position. He writes: "[Laitin's] denunciation of other people's work in his more recent scholarly phase is puzzling and hugely paradoxical. Laitin finds case studies unacceptable unless the study of a single country is transformed into a high-n research design, thereby increasing the scientific leverage...Paradoxically, Laitin's argument today amounts to denouncing his old scholarly self, so evocatively in evidence in *HC*."

I attach the paragraph (which, by the way, was not directed at a case study, but rather a methodological treatise that argued against the scientific method in the study of the social world) that contains the quoted passage:

A final example: comparativists who do qualitative case studies have no claim to disciplinary recognition by virtue of the fact that examination of a single case is a time-honored procedure in their field. Theoretical work going back to Eckstein sets constraints on what a particular case can show. More recent methodological work, exemplified in the text by King, Keohane and Verba, gives a road map on how a study of a single country can be transformed into a high-n research design, thereby increasing the study's scientific leverage. There can be no argument based on tradition justifying the minimization of leverage. New work in comparative politics must, if it is to gain respect in the wider discipline, adjust methodologically to take into account scientific advances. Pluralism without updating is not science. (Laitin 2003, 180)

Note well that I made no claim about the acceptability of case studies in this paragraph. I wrote that a traditionalist (that is, one that says that it is a time-honored approach) justification for case studies is no justification at all. In fact, I am a firm believer in case studies, but argue that (a) it is essential to use whatever data are available to inform the reader where the case fits (e.g., on the value of the dependent variable) in a larger distribution of cases and (b) it is useful to increase observations within a case in order to maximize leverage. And my position has not radically changed on this issue. Indeed, I sought in *HC* to maximize leverage. For example, I relied on survey evidence to demonstrate the lack of fit between religious practice and political views. I also added discussions of theoretical implications of my work for Nigeria's Northern Region and for the Yoruba region of Benin under different colonial incentives. Case studies are essential to our research programs; but the justification for them and the evidentiary material surrounding them must be sensitive to methodological advances in our field.

And this brings me to Varshney's comments on the selection of cases on the dependent variable. He doesn't quote my statement, but if he did, readers would see that it does not rule out the contributions of Sen or of Bates. Consider my treatment of post-Soviet violence in *IF*. Here Varshney accuses me of a contradiction, for in *IF* I select four cases that have had no interethnic violence and yet wrote a chapter on it, even after acknowledging the methodological problem. However, Varshney does not mention that shortly after the publication of *IF* I applied for funds from the Harry Frank Guggenheim Foundation to remedy that defect, and added two cases (Azerbaijan in regard to Nagorno Karabakh and Moldova in regard to Transnistria) to my sample. I published the results of the now six-country study in *Comparative Political Studies* (Laitin 2001). A piece of research that selects on the dependent variable can yield important results, but my argument is that the research program for which that piece was a part should address the problem of potential bias in future studies. There are few studies in social science with results more important than in Sen and Bates—but I'm sure both would agree that future confirmatory work would demand a more representative sample. Ignoring an inferential

problem in the name of time-tested approaches is no virtue.

A final point on my purposes in writing about the comparative method: Herrera writes, I think in reaction to my self-proclaimed “Leninist” ambitions that “Methodological choices cannot be dictated from without.” Her statement ignores the role of scientific community norms. In physics, no individual researcher can claim that in her methodology, calculus is unnecessary; in history, no individual can abjure methodological training in the reading of archives; in Classics, no individual can abjure the study of Greek. My papers on method have been written with a goal, surely less ambitious and less destructive than Lenin’s, to induce marginal change in our scientific norms such that narrative, formalization and statistical analysis become standard skills in our disciplinary tool kit.⁷

III. Conclusion

Despite my differences with the panelists in regard to the details discussed herein, I think the larger issue raised here—how to harness ethnographic data in a positivist research program that helps us understand such things as the relationship of cultural heterogeneity to democracy, to economic growth, and to social order, and not whether Laitin has been consistent across thirty years of writing—is what should most concern us as political scientists. The panelists have my respect for keeping these big issues on how best to carry out our mission at the core of their presentations.

Notes

¹ Chandra nicely identifies the complementarity between the inductive orientation in ethnography and the deductive orientation in game theory. For purposes of her own interest in mechanisms, however, she shows graphically how ethnography and rational choice “are simply two overlapping ways of identifying mechanisms, among many other ways.” There are, as Chandra points out, myriad ways to identify mechanisms (including armchair reflection). But if all methods had the same purpose with the same chance for success, we should be indifferent to methodological choice, and should not seek out complementarities. In contrast, I believe we should strategically cultivate complementarities because different methods have different strengths in the division of research labor.

² See Ferejohn (1991) for the distinction between thin and thick rational choice.

³ *HC* did not win any book awards, but was positively reviewed years after publication as a model for qualitative methods in King et al [aka KKV] (1994), and the second decade of its sales doubled what was sold in the first decade. So a concern for intersubjective reality (whatever that is) also has its rewards.

⁴ Several scholars have modeled suicide behavior with assumptions that actors were pursuing ends rationally. See Pape (2005), Bloom (2005), and Gambetta (2005) each of whom has relied on such assumptions to account for variations in behavior by potential suicide attackers.

⁵ See Booth (1961) for the forms of structure that we impose on narrative in literary criticism.

⁶ I accept Herrera’s criticism that I have not provided clear guidelines on whether the research community or the researcher should be the unit fulfilling the tripartite design. My minimalist position is that scholars who ignore research attempting to account for the same set of outcomes but using other approaches are violating scientific norms.

⁷ A related goal of mine, absent from the panel discussion, is for political science to avoid the fate of American linguistics, where formalism drove comparative language into desuetude. My insistence on narrative as part of the tripartite method is an answer to formalists and statisticians in our discipline who see no role for knowledge of place.

References (not cited in other symposium essays)

- Barth, Fredric. 1959. “Segmentary opposition and the theory of games. A study of Pathan organization.” *Journal of the Royal Anthropological Institute* 89, 5-22.
- Bates, Robert et al. 1998. *Analytic Narratives*. Princeton: Princeton University Press.
- Bloom, Mia. 2005. *Dying to Kill*. New York: Columbia University Press.
- Booth, Wayne. 1961. *The Rhetoric of Fiction*. Chicago: University of Chicago Press.
- Ferejohn, John. 1991. “Rationality and Interpretation: Parliamentary Elections in Early Stuart England.” In Kristen Monroe, ed. *The Economic Approach to Politics* (New York: Harper Collins)
- Gambetta, Diego, ed. 2005. *Making Sense of Suicide Missions*. Oxford: Oxford University Press.
- Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. New York: Cambridge University Press.
- Laitin, David. 1983. “The Ogaadeen Question and Changes in Somali Identity.” in V. Olorunsola and D. Rothchild, eds., *State Versus Ethnic Claims: An African Policy Dilemma* (Boulder, Colorado: Westview Press).
- Laitin, David. 2001. “Secessionist Rebellion in the Former Soviet Union.” *Comparative Political Studies* 34(8): 839-861.
- Laitin, David. 2004. “The Political Science Discipline.” In Edward Mansfield and Richard Sisson, eds. *The Evolution of Political Knowledge* (Columbus: Ohio State University Press).
- Pape, Robert. 2005. *Dying to Win*. New York: Random House.
- Vansina, Jan. 1985. *Oral tradition as History*. London: James Currey.
- Weber, Max. 1946/58. “The Social Psychology of the World Religions.” In *From Max Weber*, eds. W.W. Gerth and C. W. Mills (New York: Oxford University Press), 267-301.

Symposium: Alexander L. George and Andrew Bennett's *Case Studies and Theory Development in the Social Sciences*

Introduction

Jack S. Levy
Rutgers University
jacklevy@rci.rutgers.edu

During the last three or four decades, we have witnessed a sea change in the perceived utility of case study methods for theory development in international relations and comparative politics. In the 1960s, the standard view in the discipline was that most case studies were descriptive rather than theoretical, subjective and non-replicable, non-falsifiable, inattentive to the logic of causal inference, and not conducive to the cumulation of knowledge across historical cases. Now, we see a growing belief that case studies can, at least in principle, play a useful role, not only in historical description but also in the construction and/or testing of theories. Case study researchers have displayed an increasing level of methodological self-consciousness, an increasing interest in constructing research designs that will maximize their theoretical leverage, and a growing recognition that “methods” are as important for qualitative research as they are for quantitative research. Training institutes in quantitative methods are now matched by a training institute in qualitative research methods. In the last five years in particular we have seen an explosion of articles and books on qualitative methods.

There is no scholar more closely identified with the increasing methodological sophistication of case study analysis than Alexander George. Through his writings on the method of structured focused comparison, process tracing, contingent generalizations, policy-relevant theory, and related topics (George 1979, 1982); through his application of these methods and principles in the study of deterrence (George and Smoke 1974), coercive diplomacy (George and Simons, 1994), crisis management (George 1991), and a range of other topics; and through a generation of students who took his class at Stanford on “Case Studies and Theory Development,” George has played a significant role in shaping how scholars in international relations and comparative politics have organized and carried out their research. Now, after a long and fruitful collaboration with Andrew Bennett—who added distinctive contributions in the areas of philosophy of science, typological theory, and the interplay between case study and statistical methods, and who helped to refine and extend George’s earlier ideas—this work has been integrated under a single cover in *Case Studies and Theory Development in the Social Sciences* (George and Bennett 2005).

This book has already been adopted in many graduate seminars, not only by proponents of case study research but also by some scholars more associated with quantitative or

formal approaches to political science. *Case Studies and Theory Development in the Social Sciences* is likely to be highly influential and widely cited—and also, quite properly, will serve as a target of criticism. With the publication of this long-awaited book, it is time to reward the authors by subjecting their prized possession to critical scrutiny. With this consideration in mind, I organized a roundtable on the George and Bennett book for the 2005 meetings of the American Political Science Association. This symposium is an outgrowth of that roundtable. Its aim is to provide a forum for a series of brief commentaries on the contributions of *Case Studies and Theory Development* to the methodology of the social sciences. (Unless otherwise noted, all citations in this symposium are to this book.)

References

- George, Alexander L. 1979. “Case Studies and Theory Development.” In Paul Lauren, ed., *Diplomacy: New Approaches in Theory, History, and Policy*. (New York: Free Press), 43-68.
- George, Alexander L. 1982. “Case Studies and Theory Development.” Paper presented to the Second Annual Symposium on Information Processing in Organizations, Carnegie Mellon University, October 15-16.
- George, Alexander L., ed. 1991. *Avoiding War: Problems of Crisis Management*. Boulder, CO: Westview.
- George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.
- George, Alexander L. and William E. Simons. 1994. *The Limits of Coercive Diplomacy*. 2nd ed. Boulder: Westview.
- George, Alexander L. and Timothy McKeown. 1985. “Case Studies and Theories of Organizational Decision Making.” In Robert Coulam and Richard Smith, eds., *Advances in Information Processing in Organizations*. (Greenwich, CT: JAI Press), 43-68.
- George, Alexander L. and Richard Smoke. 1974. *Deterrence in American Foreign Policy*. New York: Columbia University Press.

Notes From a Generalist

Daniel W. Drezner
University of Chicago
ddrezner@uchicago.edu

Since Alexander George and Andrew Bennett warn scholars to highlight the inherent theoretical biases that researchers can carry into their analysis, I should confess to experiencing a warm glow of nostalgia when I first read *Case Studies and Theory Development in the Social Sciences* (2005). As a graduate student in the mid-nineties, I was fortunate enough to work as a research assistant for Alex George as he was mulling over

the first tendrils of this book. He proved to be an invaluable and generous font of knowledge about qualitative methods and research design. Even though Alex was emeritus at this point, and had no formal role in my dissertation committee, working for him vastly improved my own research program.

Case Studies and Theory Development represents the latest effort to rebut or revise the “hegemonic” approach to qualitative methodology as developed by King, Keohane and Verba (1994), commonly known as KKV. KKV apply the logic of inference and the tools of quantitative analysis to think about how to do qualitative case studies. Since its publication, a plethora of responses have been published (Van Evera 1997; Odell 2001; Brady and Collier 2004; Elman 2005).¹ George and Bennett do an excellent job of integrating these prior critiques into their own analysis of *Designing Social Inquiry*. There is far more in this book with which I agree than disagree.

Concurrence is boring, however, so I will focus on the disagreements, which fall into three broad categories. First, George and Bennett want to divorce methodological issues from theoretical approaches, but their discussion of how to do case studies reveals certain biases in favor of certain kinds of theory-building over others. Second, just as KKV committed the sin of oversimplifying how qualitative scholars approached their work, George and Bennett do the same in discussing the pitfalls of quantitative methods. Finally, what’s omitted from the book raises disturbing questions about the future of empirical work—as well as the state of the discipline.

Theory and Method

At a minimum, *Case Studies and Theory Development* makes it possible to economize on footnotes about qualitative methodology. In producing this text, George and Bennett have consolidated working papers (George 1982) and book chapters (George 1985) that previously elaborated on the ideas of process tracing and structured, focused comparison. The book is far greater than the sum of these parts, however. Part II serves as a useful guide for graduate students about how to design, select, perform, and infer from case studies. Part III provides a useful discussion of the ways in which proper qualitative techniques interact with policy implementation and the philosophy of science. George and Bennett have made explicit techniques that graduate students used to pick up only by osmosis. For that reason alone, this book should be incorporated into graduate student syllabi as soon as possible.

This book also offers some cogent analysis of the problems that exist in KKV’s approach. Two points stand out in particular. First, KKV assert that one way to increase explanatory leverage in qualitative analysis is to expand the number of observations by looking both within and without a case. As George and Bennett (13) point out, however, this underestimate the dangers of “conceptual stretching” that arise if increasing the number of observations requires applying theories to new cases or changing the measure of variables.² Second, George and Bennett are correct to point out that despite KKV’s claims that they wish to separate their methodological recommendations from any philosophy of science, their prior-

itization of causal effects over causal processes would privilege one class of theories over another. In particular, KKV’s methodology gives a leg up to “as if” theories of social science (Friedman 1953; Moe 1979).

In making this observation, George and Bennett (9) are trying to adhere to their own stated aim of talking about methodology but not theory: “We argue that theoretical arguments are for the most part separable from methodological debates and that case study methods have wide applicability.” However, a close reading of *Case Studies and Theory Development* calls this assertion into question. In particular, George and Bennett’s method of process tracing devotes considerable effort to examine the minutiae of government decision-making, to the point where the differences between policy principals and lower-level operators are compared and contrasted (103). Almost all of the case studies discussed in the book revolve around a chase of the paper trail that leads to key decision-making moments for foreign policy leaders.

The problem with this methodological emphasis is that it privileges middle-range decision-making approaches over structural theories with regard to hypothesis-testing. Whereas the bureaucratic politics paradigm (Allison 1971) takes greater interest in process-level variables, structuralist explanations like neorealism (Waltz 1979) will focus more on underlying causes, seeing processes as merely intervening variables.³ Consider, for example, the kind of geographical determinism that Jared Diamond (1998) uses in discussing why Europe colonized the rest of the world rather than vice versa. It would be exceptionally difficult to use process-tracing in a manner consistent with *Case Studies and Theory Development* to support Diamond’s hypothesis.⁴ Process tracing privileges theories that focus on tightly coupled cause-and-effect relationships over theories that depend on more macrohistorical evidence.

Excessive attention to one causal process can blind a researcher to the possibility that there may be substitutable causal processes at work—and in the process, privilege middle-range theories over more ambitious paradigms. Structural approaches will often posit substitutable causal processes through which the independent variable can affect the dependent variable (Most and Starr 1984). A great power might choose to influence a smaller state’s policies through carrots, sticks, or guns (Drezner 1999/2000). The offense/defense balance can affect the likelihood of war through multiple influences, including changes in grand strategies and leader perceptions (Van Evera 1999, 259-262). This is the flip side of George and Bennett’s emphasis on “equifinality.” The latter term refers to “the fact that different causal patterns can lead to similar outcomes” (161). If, however, a single structural factor determines all of those causal processes, then the issue of equifinality might not be as problematic for social science as George and Bennett believe.

Beware of the Straw Man

KKV committed the sin of underestimating the sophistication of qualitative methods when they wrote *Designing Social Inquiry*. However, there are moments when George and

Bennett fall into the same trap when discussing formal or quantitative methods. For example, on page 98 the authors argue that, “overly complex and precise formal models may posit decision-making heuristics that are ‘too clever by half’ or that no individual would actually utilize.” Two pages later, they correctly point out that, “In studying the outputs of a complex policymaking system, the investigator is well advised to work with a sophisticated model or set of assumptions regarding ways in which different policies are made in that system.” The implication from these quotations is a disturbing double standard. If a non-formal theory incorporates complexity into its decision-making calculus, then it is an example of theoretical richness; if a formal model does the same, then it is “overly complex and precise.”

On page 21, the authors assert that, “statistical methods can identify deviant cases that lead to new hypotheses, but in and of themselves they lack any clear means of actually identifying new hypotheses.” This is very much open to question. Consider, for example, the CIA-sponsored State Failure Task Force. That group of scholars amassed more than a thousand possible independent variables and used a variety of econometric filters (genetic algorithms, stepwise logistic regressions) and less formal procedures to isolate a few underlying triggers. The process generated surprising conclusions about the underlying causes of state collapse (Esty et al 1995, 1998).⁵ One could also make the larger point that quantitative methodologies have uncovered empirical puzzles that point to whole new arenas of theory-building. The interdemocratic peace is merely the most obvious example. Both formal modeling and quantitative methodologies can allow for causal complexity, theoretical richness, and hypothesis generation.

Methodological questions for the future

Despite these flaws, *Case Studies and Theory Development* will be an important methodological resource for scholars for quite some time. It is worth asking, therefore, where the discussion of qualitative methodology needs to go from here. After George and Bennett, there appear to be at least three discussions that need to take place. The first, and most important, is a more intensive focus on validity tests in weighing different pieces of qualitative scholarship. In discussing the search for documentary evidence, George and Bennett acknowledge that, “We have not yet found any book or a major article that provides an adequate discussion of the problems of weighing the evidentiary worth of archival materials” (104). This reflects our discipline’s inability to date in applying anything more sophisticated than intuition and common sense in trying to parse out contradictory paper trails. It is beyond the scope of *Case Studies and Theory Development* to develop a complete praxis on doing a case study—but that is a text that is crying out to be written.

Second, methodological debates in political science need to move beyond the stale debate of formal vs. quantitative vs. qualitative approaches. In the past decade or two a raft of new techniques have been developed and utilized in order to theorize and test political phenomenon. These include the emerging application of experimental techniques (Kinder and Palfrey

1993; Herrman et al. 1997), agent-based modeling (Cederman 2003) cognitive neuroscience (McDermott 2004), archaeological evidence (Cioffi-Revilla 1996), and online simulation and gaming (Van Belle 1998). Most of these approaches remain on the fringes of political science, and might look exotic compared to the meat and potatoes of statistics and case studies. Nevertheless, the intrusion of new methods will require a more expansive debate about marrying theory to method.

Contained within *Case Studies and Theory Development* is a final warning about one of the downsides of methodological sophistication. On page 35, George and Bennett conclude, “Because case studies, statistical methods, and formal modeling are all increasingly sophisticated... it is becoming less likely that a single researcher can be adept at more than one set of methods while also attaining a cutting-edge theoretical and empirical knowledge of his field.” Specialization and professionalization are a fact of life in the social sciences, but excessive specialization can have its perils (Stuntz 2006). As this review suggests, the theories and methods used by political scientists are more imbricated than is commonly realized. At a minimum, one would want political scientists to be avid consumers of multiple methods. If statisticians cannot evaluate the probative value of case studies, or vice versa, then the ability to generate a cumulative research agenda is put in jeopardy. As someone who has used all of the methods discussed in the quote, I hope that my methodological pluralism does not become a lost art.

Notes

¹ For a more KKV-friendly take, see Bates et al. 1998.

² Also left unstated in *Designing Social Inquiry* is how a qualitative researcher, in expanding the number of observations within a case, is to deal with the problem of autocorrelation that such an approach would present in the statistical realm.

³ Waltz (1979, 71-2), for example, believes that reductionist theories of foreign policy are useful for understanding “why different units behave differently despite similar placement in a system.” However, Waltz disdains breaking the state into its constitutive parts, concluding (65) that, “It is not possible to understand world politics simply by looking inside of states. If the aims, policies, and actions of states become matters of exclusive attention or even of central concern, then we are forced back to the descriptive level; and from simple descriptions no valid generalizations can logically be drawn.”

⁴ To be fair, Diamond’s (1998, chapter 3) analysis of the Battle of Cajamarca is an excellent example of process tracing.

⁵ See King and Zeng (2001) for a critique of the task force’s methodology.

References

- Allison, Graham. 1971. *Essence of Decision*. Boston: Little, Brown.
- Bates, Robert, et al. 1998. *Analytic Narratives*. Princeton: Princeton University Press.
- Brady, Henry and David Collier. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman and Littlefield.
- Cederman, Lars-Erik. 2003. “Modeling the Size of Wars: From Billiard Balls to Sandpiles.” *American Political Science Review* 97 (February): 135-150.

- Cioffi-Revilla, Claudio. 1996. "Origins and Evolution of War and Politics." *International Studies Quarterly* 40 (March): 1-22.
- Diamond, Jared. 1998. *Guns, Germs, and Steel*. New York: W.W. Norton.
- Drezner, Daniel. 1999. *The Sanctions Paradox*. Cambridge: Cambridge University Press.
- Drezner, Daniel. 1999/2000. "The Trouble with Carrots: Transaction Costs, Conflict Expectations, and Economic Inducements." *Security Studies* 9 (Autumn/Winter), 188-218.
- Elman, Colin. 2005. "Explanatory Typologies in Qualitative Studies of International Politics." *International Organization* 59 (Spring): 293-326.
- Esty, Daniel et al. 1995. *Working Papers: State Failure Task Force Report*. McLean, VA: Science Applications International Corporation.
- Esty, Daniel et al. 1998. *The State Failure Task Force Report: Phase II Findings*. McLean, VA: Science Applications International Corporation.
- Friedman, Milton. 1953. "The Methodology of Positive Economics," in Milton Friedman, *Essays in Positive Economics* (Chicago: University of Chicago Press).
- George, Alexander. 1982. "Case Studies and Theory Development." Working paper, Stanford University.
- George, Alexander and Timothy McKeown. 1985. "Case Studies and Theories of Organizational Decision-Making." In R. Coulam and R. Smith, eds. *Advances in Information Processing in Organizations*. (Greenwich, CT: JAI Press).
- Herrman, Richard, et al. 1997. "Images in International Relations: An Experimental Test of Cognitive Schemata." *International Studies Quarterly* 41 (September), 403-433.
- Kinder, Donald and Thomas Palfrey. 1993. "On Behalf of an Experimental Political Science." In Kinder and Palfrey, eds., *Experimental Foundations of Political Science* (Ann Arbor: University of Michigan Press).
- King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton: Princeton University Press.
- King, Gary and Langche Zeng. 2001. "Improving Forecasts of State Failure." *World Politics* 53 (July), 623-58.
- McDermott, Rose. 2004. "The Feeling of Rationality: The Meaning of Neuroscientific Advances for Political Science." *Perspectives on Politics* 2 (December), 691-706.
- Moe, Terry. 1979. "On the Status of Scientific Models." *American Journal of Political Science* 23 (February), 215-243.
- Most, Benjamin, and Harvey Starr. 1984. "International Relations Theory, Foreign Policy Substitutability, and 'Nice' Laws." *World Politics* 36 (1984), 383-406.
- Odell, John. 2001. "Case Study Methods in International Political Economy." *International Studies Perspectives* 2 (May): 161-176.
- Ragin, Charles. 1987. *The Comparative Method*. Berkeley: University of California Press.
- Stuntz, William. 2006. "Future Shock." *The New Republic*, (27 February).
- Van Belle, Douglas. 1998. "Balance of Power and System Stability: Simulating Complex Anarchical Environments over the Internet," *Political Research Quarterly* 51 (March), 265-282.
- Van Evera, Stephen. 1997. *Guide to Methods for Students of Political Science*. Ithaca: Cornell University Press.
- Van Evera, Stephen. 1999. *Causes of War*. Ithaca: Cornell University Press.
- Waltz, Kenneth. 1979. *Theory of International Politics*. New York: McGraw-Hill.

A Major Milestone with One Major Limitation

John S. Odell

University of Southern California
odell@usc.edu

Alexander George is one of political science's most respected innovators and teachers of qualitative research design, and he and his former student Andrew Bennett have now teamed up to produce what is likely to become an influential text. It presents George's influential methodological ideas together in one place for the first time and develops and defends them further with new contributions by Bennett added through close collaboration. I believe their book makes several valuable new contributions and also has a few limitations, one of them rather important.

Major Contributions

One major contribution is strong advocacy on behalf of qualitative case study (CS) methods, especially process-tracing. Doctoral education in political science was biased in favor of quantitative methods for decades (until the 1990s, at least), and this imbalance surely diminished our profession's receptivity to qualitative methods. This new volume adds a welcome corrective, joining King, Keohane, and Verba (1994) and more recently, Brady and Collier (2004) and Sprinz and Wolinsky-Nahmias (2004). Chapter 1 emphasizes several comparative advantages of CS methods for theory development, especially theory that will help practitioners in governments. Perhaps the most significant strength is the fertility of CS research in generating new concepts and hypotheses that are likely to be empirically valid. The book argues that case studies are also especially strong for improving the validity of our concepts and avoiding "conceptual stretching"; for exploring the operation of causal mechanisms; and as ways to accommodate complex causal relations such as equifinality, interaction effects, and path dependency. I would add that case studies often have an advantage over statistical methods when the objective is to study a process. Case studies also add to the literature more information about the cases studied than is possible with statistical methods covering the same cases, information which may be used by later researchers. Bennett and George also defend against widely-heard criticisms of CS methods—such as the selection bias problem—that fail to give sufficient credit for their advantages.

Second, the book offers a philosophical standard for evaluating research and methods different from the standard underlying statistical explanation, strengthening its defense against critiques from the latter quarter. Chapter 7 draws especially on Wesley Salmon's innovations in philosophy of science (1990), defining theoretical explanation as knowledge of the hidden causal mechanisms that generate observable events. George and Bennett contrast exploration of causal mechanisms with the deductive-nomological model and the estimation of "causal

effects” as defined by King, Keohane, and Verba (1994). The book includes criticism of *Designing Social Inquiry* at several points where they feel this influential text got it wrong. They also reject the postmodern position that it is impossible to separate the object from the researcher or to develop cumulative theory. In developing a philosophical foundation the book goes deeper than the authors’ earlier publications.

At the same time, they are not satisfied with the quality of some published case studies. In hopes of improving scholarly practice, they offer a practical handbook for students, laying out steps by which to implement the techniques pioneered by George with collaborators Richard Smoke and Timothy McKeown. The main ideas had already been published, but presenting them together in this book allows greater development and more examples.

The book details several varieties of within-case process-tracing, defining the idea flexibly—it ranges from atheoretical narrative at one extreme to even a variant that “constructs a general explanation rather than a detailed tracing of a causal process” (211). George and Bennett argue that process-tracing is especially useful for finding multiple causal paths that lead to the same outcome, for explaining deviant cases, for improving theories that are not specified sufficiently, and for testing theories. They emphasize that adding within-case techniques can improve causal inferences in a study that compares multiple cases.

Chapter 11 elaborates on one of their most distinctive ideas—typological theorizing. The basic idea is to develop a typology of alternative causal pathways or sets of variables that produce a phenomenon of interest (states’ contributions to an alliance, in their example). The typology can be developed either inductively by studying one case at a time, or deductively by first defining a logically complete space of potential causal variables, then reducing the space and selecting cases to study each type. The technique is illustrated concretely by reference to Bennett, Lepage and Unger (1997) on burden sharing during the first Persian Gulf war. George and Bennett find that the process of attempting to construct theories typologically has a salutary effect on the analytical imagination. It shifts attention from univariate thinking toward thinking about combinations of variables and multiple causal paths, and it exposes inconsistencies in earlier theorizing. Although this technique has not been widely adopted thus far, it is interesting and this book might encourage more attempts.

The book includes a chapter reviewing the literature on the inter-democratic peace and shows the complementarity of case study, statistical, and formal methods over time as well as the possibility of cumulative theory development. A long appendix also describes a number of publications in detail to illustrate the variety of ways in which within-case causal analysis has been used or approximated.

Some Limitations

One minor limitation is that while the proposed methods are general, the book’s codification of best practice is limited largely to illustrations from international security studies, supplemented by some references from comparative politics.

The social scientist will need other books to learn how case studies have been used extensively in research on the international political economy, environmental politics, human rights, institutional change, negotiation, domestic government program evaluation, educational reform, life in primitive tribes and urban slums, and other domains of social science (see Sprinz and Wolinsky-Nahmias [2004] for fuller coverage of international relations).

Second, *Case Studies and Theory Development in the Social Sciences* tends to equate case study methods with the historian’s craft and with a particular type of evidence—archives and other documents. This type is certainly important and the authors give wise warnings about the pitfalls of using archives. But most social science methodologists define case study methods more broadly. Many in international relations and other specialties use interviews to collect evidence before archives are open. Other case studies are based on direct observation, participant observation, or statistical data. Some case studies are largely quantitative. Other methods texts (e.g., Yin 1994) provide more comprehensive coverage and advice in this respect.

More important, a neutral reader might feel that the inherent drawbacks of CS methods are not presented quite as forcefully as the advantages. The book does raise and respond to complaints and advocates methodological pluralism. To take an example, one basic complaint is that case study methods do not support generalizations well because the cases chosen may not be representative of the universe covered by the theory. The book recommends that case study researchers concentrate on a subclass of the event of interest, to find conditions under which specified outcomes occur, and to make clear that one is not claiming to generalize beyond that subclass in a given study (31-32). But this response seems only to transfer the problem to the subclass level rather than resolve it. Typically the case study cannot show that the case selected is representative of events in the theoretical subclass either.

To take another related issue, the authors argue that CS methods are useful for testing theories, while for many scholars the main point would be that testing is where all case methods are inferior to statistical methods. The notion of testing has more than one dimension and commentators may disagree because they emphasize different dimensions. A test is more rigorous, first, the more observations the theory has faced and the more representative the sample. A second dimension is how precise the measurement is; the looser a concept’s effective meaning, the easier it is to fit it to more facts. Statistical methods are inherently superior on both these dimensions to a method that can observe only a small number of cases whose representativeness is unknown (except regarding a deterministic theory) and a method that does not quantify its measurements. A third facet of testing, however, is how many rival explanations have been eliminated by the analysis. Here the case study is able to consider more candidates, including what is often called the context, more thoroughly than the statistical method, in the cases studied by each. The case study can also check in finer detail for mechanisms that theories assume con

nect cause to effect. This strength of case methods mitigates their overall disadvantage when it comes to testing. There are also ways to make a case study more rigorous in all three respects (Odell 2004). But on balance my advice to students is in most cases to avoid claiming to have “tested” a theory using only case studies, and instead to make more nuanced claims of other valuable contributions.

I believe the book’s most important shortcoming, though, is that in its enthusiasm to promote process-tracing, it goes too far in denigrating a valuable alternative. The book strikes me as biased against cross-case comparative methods. Chapter 8 is the place where it concentrates on comparative methods, beginning rightly with Mill’s foundational methods of agreement, difference, and concomitant variation that represent the logic of experimentation. I believe the book understates the accomplishments and exaggerates the problems of what it calls controlled comparison—and elsewhere the book downplays the shortcomings of process-tracing.

Chapter 8 is almost completely silent about the advantages of comparative methods (to be discussed in a moment). It jumps right to criticisms and limitations. Even though eminent comparativists have used and recommended Mill’s methods, the book does not analyze exemplars to teach students what best practice looks like—something it does copiously for process-tracing. This chapter mentions Skocpol’s *States and Social Revolutions* (1979) but mainly to illustrate how process-tracing can address limitations of controlled comparison. To take another example, their appendix mentions *Double-Edged Diplomacy* (Evans et al. 1993), but surprisingly they never mention that every chapter of that book was a two-case controlled comparison. Outside chapter 8, George and Bennett mention some other comparative studies approvingly, but for some reason they do not report the accomplishments of the comparative method in this chapter.

Chapter 8 complains that an inherent limitation of controlled comparison is that another case may be discovered later that does not confirm the argument (155). But this is true of any case study, and yet chapter 10 on process-tracing does not mention this weakness. That chapter does warn that more than one explanation may be consistent with evidence but it praises the technique for being able to make a partial contribution (222). The same could be said equally for comparative methods but this book does not say so; on the contrary it uses this limitation to direct attention away from comparative methods.

It is true that Mill himself said his methods were difficult to use in the study of societies, given his scientific goals. Bennett and George put it this way: “Mill’s methods can work well in identifying underlying causal relations only under three demanding assumptions. First, the causal relation being investigated must be a deterministic regularity involving only one condition that is either necessary or sufficient for a specified outcome. Second, all causally relevant variables must be identified prior to the analysis...Third, cases that represent the full range of all logically and socially possible causal paths must be available for study” (155). Their main conclusion is that “unfortunately, practically all efforts to make use of the

controlled comparison method fail to achieve its strict requirements.”...“Researchers urgently need an alternative to the experimental paradigm” (151).

This dispute between distinguished users and critics of these methods might be sorted out better if we focus on the scientific goal that underlies the methods being advocated. J.S. Mill in 1843 was writing a book on logic, and the famous methods appear in his Book III on Induction. He defines induction as “the operation of discovering and proving general propositions” (Mill 1843, 208). Mill believed the goal of science was to “prove” the truth of “invariable laws” “with certainty” (e.g., 278, 282). He uses these terms repeatedly. It is with this ultra-ambitious goal in mind that he defines his method of difference, for example, saying “the two instances which are to be compared with one another must be exactly similar, in all circumstances except the one which we are attempting to investigate” (281).

Thus the claim that Mill’s methods require those hyper-demanding assumptions is valid only if the researcher aims to prove with 100% certainty that a general law is true invariably, by observing only two or three cases. Of course their own method of process-tracing in the single case would also have to be discarded if this were the only standard. In the vast majority of applications it too by itself fails to verify with airtight certainty that a general law is true invariably, i.e., in the many cases that have not been observed. But the point is that most political scientists abandoned this nineteenth century version of positivism long ago. After the heyday of Victorian optimism, verificationism gave way to falsificationism. Then Lakatos falsified Popper (much to Popper’s disappointment). Probabilistic explanation came to be the generally accepted mode of causal analysis.

A more even-handed assessment would treat controlled comparison using a bit of the flexibility shown when defining process-tracing. It is more productive in practice to think of controlled case comparison as an approximation (not an exact match) to the logic of experiment, with the goal of making at least a significant partial contribution to (not to prove or test) a probabilistic causal explanation, in most cases. (Any more decisive conclusions achieved by testing the occasional deterministic theory would be further to its credit.) Today’s methods of difference and concomitant variation can be defined roughly as research designs in which two or more cases differ on a supposed causal variable C and on a dependent variable E and are *similar* on at least *several* other relevant causal variables. If the difference is in the expected direction, then we may conclude that the design provides significant support for the inference that C influenced E, and stronger support than without the comparison and without the matching. Support is a matter of degrees, not a dichotomy, in most research. I believe this is what many contemporary scholars meant when they claimed to be using Mill’s methods.

This revised version of Mill delivers powerful, well-known advantages. Contrasting cases carefully adds confirmation that comes from observing differences in the cause that cannot be observed in many single case studies. The deliberate selection of cases that match in other relevant respects elimi-

nates threats to the inference's validity that would often be present without this design. A multiple-case study designed this way will be more convincing than many single-case studies and many studies of multiple cases selected without regard to theory. At the same time, the comparative case method is able to deliver the payoffs of the case study that are lacking in statistical studies. It is an intermediate technique. And a sequence of comparative studies can achieve more than any one alone.

This "approximate" method of difference does require caveats. History always presents, in a given set of cases, variables besides the study variable that also varied and also might have contributed to the outcome variation. The author must therefore acknowledge that more research will be needed to sort out those causes with greater clarity and certainty. And the more such variables that are present in a given set of cases, the weaker the study's design will be. It will usually be too much to claim, from only two or three cases, to have found all possible causal pathways or to have ruled out all interaction effects. We do run a risk of being misunderstood if we claim to be using Mill's methods but do not specify that we aspire to a more realistic philosophical standard than Mill's own. But Copi and Cohen (1990, 407), a book George and Bennett cite approvingly, sums up the conclusion well: "Mill's methods are more limited instruments than Bacon and Mill conceived them to be, but within those limits they are indispensable."

Every other social science method also turns out to be inherently imperfect in some respect and also requires caveats. Every process-tracing study requires the caveat that the case could be unrepresentative, even of the type being theorized. A regression study with 10,000 observations will be misleading if measurement error or omitted variables are significant, and many do not demonstrate directly that the alleged causal mechanism actually operates. Experiments have their own inherent limitations. Unfortunately there simply is no escape from methodological tradeoffs and complementarity in social science.

Thus, I believe additional materials will be needed in my course on qualitative research design to teach students how to use comparative methods well. But overall, this book is the fullest presentation of the methodological advice of one of our most admired leaders. Its many strengths, due to long experience and sophisticated reflection by Bennett as well, are likely to earn it wide and lasting influence in political science. It will help scholars raise the quality of future case study research.

References

- Bennett, Andrew, Joseph Lepgold, and Danny Unger, eds. 1997. *Friends in Need: Burden-Sharing in the Persian Gulf War*. New York: St. Martin's Press.
- Brady, Henry E., and David Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. New York: Rowman and Littlefield Publishers.
- Copi, Irving M., and Carl Cohen. 1990. *Introduction to Logic*, 8th ed. New York: Macmillan.
- Evans, Peter, Harold Jacobson, and Robert Putnam, eds. 1993. *Double-*

- Edged Diplomacy*. Berkeley: University of California Press.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton: Princeton University Press.
- Mill, John Stuart. [1843] 1900. *A System of Logic*, 8th ed. New York and London: Harper and Brothers.
- Odell, John S. 2001. "Case Study Methods in International Political Economy." *International Studies Perspectives* 2:2, 161-76. Reprinted and revised in Detlef F. Sprinz and Yael Wolinsky-Nahmias, eds. 2004. *Models, Numbers and Cases* (Ann Arbor: University of Michigan Press), 56-80.
- Salmon, Wesley. 1990. *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.
- Skocpol, Theda. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.
- Sprinz, Detlef F. and Yael Wolinsky-Nahmias, eds. 2004. *Models, Numbers and Cases: Methods for Studying International Relations*. Ann Arbor: University of Michigan Press.
- Yin, Robert K. 1994. *Case Study Research: Design and Methods*, 2nd ed. Thousand Oaks, California: SAGE Publications.

Well Worth the Wait

Jack S. Levy
Rutgers University
jacklevy@rci.rutgers.edu

In the early 1980s, when I wanted to provide some guidance to graduate students on case study methodology, I would direct them to Alexander George's published 1979 essay and then to his 1982 conference paper on "Case Studies and Theory Development." I continued to assign the unpublished essay for nearly fifteen years, until early draft chapters of the George and Bennett book manuscript became available. With the publication of George and Bennett's *Case Studies and Theory Development in the Social Sciences*, any uncertainty about which primary text to use for a course on qualitative research methods has evaporated, even with the recent proliferation of many excellent articles and books on the case study method.¹

This long-awaited book has been well worth the wait. It refines George's pathbreaking work on the method of structured focused comparison, process tracing, and other topics in case study methodology; makes important advances in his earlier ideas on typological theory; thoroughly grounds the approach in the philosophy of science; argues for the compatibility of case study methods and quantitative methods; and provides a useful "how to" manual for students. It will be widely read and widely cited. For any graduate student using case studies in his or her dissertation, or for anyone interested in a minimal familiarity with the many diverse methodological approaches in the discipline,² this book is like the American Express card. You should not leave school without it.

There are many things that I like about the book. At the most basic level, I like the theme of converting descriptive explanations of historical cases into analytic explanations, a theme that George (1979, 1982) has been emphasizing for a quarter century. I like the emphasis on the interplay between theory and evidence, a theme that goes back at least to George and Smoke (1974). Theory is used to structure the interpreta-

tion of the case; evidence from the case is used to refine or recast the theory; and the revised theory is used to guide further empirical applications, either on new cases or on different aspects of same case. This is reminiscent of Popper's (1962) notion of conjectures and refutations, and it reflects a fairly common-sense notion of scientific progress. What is new here is that the interaction between theory and evidence is both more deeply grounded normative conceptions of scientific progress (Lakatos 1970) and linked to key issues of research design.

I also like George and Bennett's distinction between epistemological logic and methodological logic. They argue that King, Keohane, and Verba's (1994) argument about a single logic of inference applies to the level of epistemology but not to the level of methodology;³ that the similar epistemological logics underlying statistical, experimental, and case study methods mean that these methods are complementary rather than conflicting; that the different methodological logics provide each of the different approaches with distinctive advantages and disadvantages; and that a multi-method research design can use the advantages of one method to compensate for the disadvantages of another.

George and Bennett's extended discussion of case selection and alternative case selection designs is another major strength of the book. While they accept the argument, associated with King, Keohane, and Verba (1994), that we should derive many testable implications of our theories, George and Bennett argue that for various aspects of theory development, including theory testing, not all cases are equally important. They use Bayesian logic to argue that some cases provide more leverage than others for illuminating or testing particular theories, and they emphasize the utility of most likely and least likely research designs. Testing a theory on a case where our theoretical priors suggest that the theory is unlikely to be validated—either because the values of many of the theory's key variables point in the other direction or because the theory's scope conditions are not fully satisfied—can provide a great deal of leverage for increasing our confidence in the validity of the theory, if the data supports the theory. Theoretical leverage from such a “least likely” case is further enhanced if that case is an “easy” (most likely) one for the leading alternative theory (in terms of theoretical priors) and if the data do not support that theory. Similarly, if one's theoretical priors suggest that a theory is highly likely to be confirmed, and if the data do not support the theory, that result can be quite damaging to a theory.⁴ It is important to note that the logic of inference in these designs is quite asymmetric. While evidentiary support for a theory from a least likely case or lack of support from a most likely case provides substantial theoretical leverage, evidentiary support for a theory from a most likely case or lack of support from a least likely case provides relatively little basis for generalizing.

Most likely and least likely case designs are hardly new in the social sciences (Eckstein 1975). George and Bennett's contribution is to explicitly ground the analysis in Bayesian logic; to incorporate both process tracing and within-case comparisons; to link the logic underlying these designs to the logic of

three cornered tests (Lakatos 1970), by emphasizing not only the prior likelihood of the theory and its evidentiary support but also the likelihood of and support for alternative theories; and to tie the discussion to the utility of single case studies for generalizing beyond the data in the particular case.⁵

Another positive feature of *Case Studies and Theory Development in the Social Sciences* is its argument that case study methods, including within-case process tracing as well as cross-case methods, are compatible with nearly *any* theoretical orientation, including rational choice (208). This view is the core of the “analytic narrative” research program (Bates et al 1998), and it has long struck me as obvious.⁶ In terms of the sociology of the profession, however, I think that one of the reasons some scholars have embraced the extraordinary rise of case study methodology, at least in political science, is the belief that those methods gave them a useful weapon in theoretical debates pitting decision-making or cultural approaches against rational choice. In the context of these paradigmatic battles, George and Bennett provide a useful reminder that case study methods are not tied to a particular theoretical orientation.⁷ We should remember that George's classic analysis of deterrence failure (George and Smoke 1974), which was the first systematic application of the method of structured focused comparison and which may have done more than any other piece of research to convince international relations scholars of the theoretical value of systematic controlled comparison and process tracing, was primarily rationalist in its theoretical orientation.

Just as case study methods can be useful for the development and perhaps testing of a wide range of different theories, theories can be tested by different methodologies. I agree strongly with George and Bennett's emphasis on multi-method research. They argue that case study, statistical, and formal methods each have distinctive strengths and weaknesses, and that the combination of several methods can give the researcher additional leverage and help compensate for the limitations of each method. They show how the application of multiple methods to the question of the democratic peace has provided a much deeper understanding of that phenomenon than would have been possible through any single method. The same is true for other research programs in the discipline.⁸

In emphasizing the benefits of multi-method research, George and Bennett (34-35) recognize that as methods have become more and more complex and more difficult and time consuming to learn, it is increasingly unlikely that a single researcher will be able to develop the requisite methodological skills in several distinct areas while staying on top of the theoretical and empirical developments in his or her field. They emphasize the importance of individual scholars becoming expert in one method and at least “conversant” with others, suggest that this will facilitate collaborative research, and argue that collaborative research among scholars with different comparative advantages but with an ability to understand and communicate with each other is particularly valuable for the advancement of knowledge. I strongly endorse this argument.

Although George and Bennett's advocacy of multimethod research is welcome, I think they could have gone further in suggesting how different approaches might be combined to develop better theories in the social sciences. Let me suggest one example. With so many applications of process tracing to decision-theoretic models, George and Bennett could have said more about how case study and experimental methods might be combined. In the study of decision making the perceptions, expectations, and choices of actors are among the critical variables, yet they are difficult to study empirically because of limited access to the relevant data. Supplementing case studies with experimental approaches might give us greater confidence in the validity of our findings.

One possibility is illustrated by Mintz et al's (1997) application of a computerized "decision board" methodology to explore the sequence of individual thought processes in judgment and decision-making.⁹ This is a different kind of process tracing, with its own distinctive set of advantages and disadvantages. The decision board methodology provides greater confidence that one has tapped the thought processes of any particular individual, but it raises questions as to whether the subjects of laboratory experiments are comparable to political decision-makers, and whether the experiment can simulate the stakes and psychological environment confronting political leaders when they make decisions. A careful combination of process tracing through case studies and process tracing through a decision board approach might be particularly useful in the exploration of some theoretical questions.

I would also like to address a common argument, found in many discussions of the benefits of case study methods or in particular case study applications, about the unique advantages of case study methods, process tracing in particular. That is the argument that statistical studies are useful for the exploration of correlations between variables but not for the analysis of causal mechanisms, while process tracing through case studies has an absolute advantage in uncovering causality. Some state this "causality gap" more strongly than others, and insist that statistical methods simply cannot get at causality at all. George and Bennett strike a reasonable balance. After a discussion in which they emphasize the role of process tracing in identifying intervening causal processes, and indeed suggest that it has some distinct advantages for those purposes, they argue that "While process tracing can contribute to theory development and theory testing in ways that statistical analysis cannot (or can only with great difficulty), the two methods are *not* competitive. The two methods provide different and complementary bases for causal inference..." (206-08).

I do not object to George and Bennett's view on this issue, but let me take this opportunity to address the stronger form of the argument, the common idea that statistical methods cannot get at intervening causal mechanisms. Causal mechanisms are best treated as unobserved theoretical constructs (George and Bennett 135-37), which are posited by a theory or hypothesis to help provide a logically complete explanation. Hypothesized causal mechanisms generate observable implications, and we test the hypothesized causal mecha-

nism by testing its observable implications. Many hypothesized causal mechanisms imply that particular actors will hold particular beliefs or take particular actions at particular times. These predictions can often be validated through process tracing in intensive cases studies. Causal mechanisms also generate implications that predict correlations between variables. Such implications, and hence the causal mechanisms that generate them, can sometimes be tested through statistical methods.

One good example of the use of statistical methods to test hypothesized causal mechanisms is Schultz's (1999) examination of two alternative models of the democratic peace and of the foreign policies of democracies more generally—the signaling mechanism based on revelation of information, and the institutional constraint mechanism. He formalized each theoretical argument, derived predictions from each model, identified where those respective predictions are congruent and where they are contradictory, selected cases based on the later, conducted statistical tests, and concluded that the signaling mechanism was more consistent with the data. Schultz's tests focus on outcomes rather than processes, but those outcomes have direct implications for processes and causal mechanisms that drive them. Schultz's statistical analysis increases our confidence in the validity of his signaling model of the democratic peace. That confidence is further enhanced by similar findings from Schultz's case studies, just as confidence in a process tracing analysis of causal mechanisms in one case would be reinforced by statistical analysis of traces of those mechanisms in a larger number of cases.

It is undoubtedly true that *Case Studies and Theory Development in the Social Sciences* will have a positive impact on the conduct of case study research, by emphasizing its theoretical orientation, by suggesting ways of maximizing theoretical leverage while recognizing potential pitfalls, and more generally by systemizing a method that some enlightened practitioners have been following for years. The book will also have a broader impact on the discipline of political science, by emphasizing the compatibility of different methods and making a strong argument for multimethod research, and more generally by enhancing the methodological self-consciousness of scholars across various research communities. It was well worth the wait.

Notes

¹ In my spring 2006 seminar on Qualitative Research Methods, my required books included Ragin (1987), King, Keohane, and Verba (1994), Brady and Collier (2004), and George and Bennett (2005).

² This should include everyone.

³ Brady and Collier (2004) emphasize a similar theme, as reflected in the subtitle of their book: "Diverse Tools, Shared Standards."

⁴ The logic of least likely case design is based on what I call the Sinatra inference—if I can make it there I can make it anywhere. Similarly, the logic of most likely case design is based on the inverse Sinatra inference—if I cannot make it there, I cannot make it anywhere (Levy 2002).

⁵ Eckstein (1975) also emphasized the theoretical utility of individual cases. See also Rogowski (1995). For an argument that case studies can contribute to theory construction but not to theory test-

ing, see Achen and Snidal (1989).

⁶ For a similar view from sociology see Kiser (1996).

⁷ In fact, George and Bennett's emphasis on the importance of middle range theory is an implicit acknowledgment of their belief that paradigmatic debates are not the most useful way of advancing knowledge.

⁸ One in international relations is the diversionary theory of war.

⁹ For an application of decision boards to the study of judgment in voting, see Lau and Redlawsk (1997).

References

- Achen, Christopher H. and Duncan Snidal. 1989. "Rational Deterrence Theory and Comparative Case Studies." *World Politics* 41:2 (January), 143-69.
- Bates, Robert H., Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry Weingast. 1998. *Analytic Narratives*. Princeton: Princeton University Press.
- Brady, Henry E. and David Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman & Littlefield.
- Eckstein, Harry. 1975. "Case Studies and Theory in Political Science." In Fred Greenstein and Nelson Polsby, eds., *Handbook of Political Science*, vol. 7. (Reading, MA: Addison-Wesley), 79-138.
- George, Alexander L. 1979. "Case Studies and Theory Development." In Paul Lauren, ed., *Diplomacy: New Approaches in Theory, History, and Policy* (New York: Free Press), 43-68.
- George, Alexander L. 1982. "Case Studies and Theory Development." Paper presented to the Second Annual Symposium on Information Processing in Organizations, Carnegie Mellon University, October 15-16.
- George, Alexander L. and Timothy J. McKeown. 1985. "Case Studies and Theories of Organizational Decision Making." *Advances in Information Processing in Organizations* 2, 21-58.
- King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton University Press.
- Kiser, Edgar. 1996. "The Revival of Narrative in Historical Sociology: What Rational Choice Theory Can Contribute." *Politics & Society* 24 (September), 258.
- Lakatos, Imre. 1970. "Falsification and the Methodology of Scientific Research Programs." In Imre Lakatos and A. Musgrave eds., *Criticism and the Growth of Knowledge* (New York: Cambridge University Press), 91-196.
- Lau, Richard R. and David P. Redlawsk. 1997. "Voting Correctly." *American Political Science Review* 91:3 (September), 585-598.
- Levy, Jack S. 2002. "Qualitative Methods in International Relations." In Michael Brecher and Frank P. Harvey, eds., *Millennial Reflections on International Studies* (Ann Arbor: University of Michigan Press), 432-54.
- Mintz, Alex, Nehemia Geva, Steven B. Redd, and Amy Carnes. 1997. "The Effect of Dynamic and Static Choice Sets on Political Decision Making: An Analysis Using the Decision Board Platform." *American Political Science Review* 91:3 (September), 553-66.
- Popper, Karl R. 1962. *Conjectures and Refutations*. New York: Basic Books.
- Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.
- Rogowski, Ronald. 1995. "The Role of Theory and Anomaly in Social-Scientific Inference." *American Political Science Review* 89:2 (June), 467-70.
- Schultz, Kenneth A. 1999. "Do Democratic Institutions Constrain or Inform? Contrasting Two Perspectives on Democracy and War." *International Organization* 53:2 (Spring), 233-66.

Case Studies and the Philosophy of Science

David Dessler

College of William and Mary
dadeds@wm.edu

From the perspective of the philosophy of science, the main weakness of standard books on research methods in the social sciences is that they neglect or suppress discussion of the philosophical underpinnings of science. To some readers this might seem a virtue—after all, is not the purpose of these books to explain method, rather than debate philosophy?—but the fact is that every social scientist, in committing to a particular method of analysis, takes sides in important philosophical disputes. Indeed, debates over method in social science are typically at least in part debates over issues in the philosophy of science. For better or worse, the philosophy of science is unavoidable in discussions of social science methodology.

From this perspective, *Case Studies and Theory Development* is a breakthrough work. Like other methods books, it presents "how to" arguments concerning proper procedures of social science research. Unlike those other books, it expounds systematically the philosophical underpinnings of these arguments. The authors do not debate philosophy for philosophy's sake, but rather apply the results and insights of philosophy to practical concerns in the methodology of social science.

In three areas George and Bennett make particularly important philosophical contributions. The first concerns the nature of causal explanation. The second deals with the nature explanation in history. And the third concerns our understanding of the real-world relevance of social research.

One of the most important objectives in any science is the causal explanation of individual events. We want not only to develop general theory; we also aim for explanatory understanding of concrete events. What explains the French Revolution? What caused World War I? Why was democracy stable in the Netherlands between 1917 and 1967? Why was the Versailles Treaty rejected by the U.S. Senate? The answers to such questions, George and Bennett argue, rest on the identification of relevant *causal mechanisms*. Here the authors target "the traditional positivist model" of "deductive-nomological" (D-N) explanation, according to which an event or outcome is explained by showing it is an instance of a general law. In their view, the D-N model suffers from a major flaw: it cannot distinguish between causal laws and accidental regularities. Explanation by mechanisms, in contrast to the D-N model, explains by detailing the processes connecting cause and effect in particular cases. In George and Bennett's view, "causal mechanisms provide more detailed and in a sense more fundamental explanations than general laws do" (141).

George and Bennett's analysis of causal mechanisms is

cogent and sophisticated. They are persuasive in arguing that explanation, to be successful, requires the identification of one or more mechanisms. However, one may wonder whether they have dismissed the D-N model too quickly. For it is not clear that explanation by laws and explanation by mechanisms are necessarily *alternatives*. Laws identify regularities, and mechanisms account for these regularities by detailing the processes that make them recur. Is it not possible that explanation depends on *both* laws and mechanisms? A law shows that the event in question was something to be expected, and the associated mechanism provides an understanding of this expectation. Both components are needed for a successful explanation. For example, to borrow an illustration from George and Bennett (132), we might explain rainy weather by the passage of a cold front. This is a good explanation because the passing of cold fronts in general is associated with rainy weather (a law), and we account for this regularity by detailing the processes through which the cold front lifts warm moist air and cools it so that a storm is generated (the mechanism). Note that by identifying the associated mechanism, we repair the main deficiency George and Bennett identify in the D-N model: we show the law to be causal rather than a mere accidental association.

George and Bennett's account of historical explanation seems to confirm the need for laws as well as mechanisms in explanation. In taking up the question of the relationship between process-tracing and history, the authors defend a type of explanation that positivist philosophers labeled "genetic." In a genetic explanation, a phenomenon is explained by showing it is the endpoint of a sequence of events that together form a pathway or trajectory through space and time. At each point in this sequence, George and Bennett note, the analyst is responsible for adducing an appropriate "covering law," that is, a law that covers that particular segment of the overall pathway (227-8). This is precisely what positivist philosophers such as Carl Hempel argued, and it confirms the relevance of laws to explanation even in cases where explanation is essentially historical in character.

The authors confront directly the question of "whether process-tracing is similar to historical explanation and whether process-tracing is anything more than 'good historical explanation'" (224). They respond, reasonably enough, that the answer to this question depends on just how one conceptualizes historical explanation. One type of such explanation, of course, is "a detailed narrative or story presented in the form of a chronicle that purports to throw light on how an event came about" (210). They present such narratives or stories as merely first steps toward more systematic and careful process-tracing that would be relevant for a theoretically oriented explanation. But one might ask here whether George and Bennett do not underestimate the power of process-tracing in even relatively atheoretical narratives. Here it is useful to distinguish between general and singular causation. Examples of general causation include, "Smoking causes cancer," and "Bipolarity promotes system stability." Examples of singular causation include "Smoking caused Fred's cancer," and "The postwar system was stable because it was bipolar."

Questions about the general workings of particular causal factors (Does a fish diet promote cardiovascular health? Do military organizations favor offensive doctrines? How do revolutions affect the propensity to war?) invite answers that describe general causation. Questions as to what factors produced certain outcomes in particular (Why did Joe get sick? What explains French military doctrine between the world wars? What brought about the Napoleonic wars?) call for the identification of singular causes.

George and Bennett, focused as they are on the development of theory in social science, discuss process-tracing in terms of its contribution to our knowledge of general causation. But process-tracing is equally important to the localization of singular causes in historical explanation. An excellent example in the study of world politics is given in David Holloway's *Stalin and the Bomb*. Holloway shows that nuclear weapons caused the Soviets to be more cautious and at the same time more stubborn in their dealings with the West. He makes no attempt to determine the adequacy of claims about the effects of nuclear weapons on the foreign policies of states in general. The latter might be useful knowledge in its own right, but the historical analysis of the Soviet case stands on its own, buttressed by careful process-tracing that follows the logic outlined by George and Bennett.

A third area in which George and Bennett make an original and important philosophical contribution concerns the relation between theory and practice. This is a topic almost entirely ignored in standard methods books, which typically urge social scientists to carry out research that addresses important real-world problems, but leave the meaning of this injunction unclear. This has put political scientists in the position of having to defend themselves against the charge that their research is "method-driven" rather than motivated by its practical relevance. George and Bennett tackle the issue of relevance directly, defining real-world importance in terms of the challenges facing decision-makers in government. Here they are critical of highly general theories in social science, as these accounts "offer little insight into how decision-makers can choose policy instruments to influence outcomes." They favor instead "middle-range theories" that "provide better guidance about when various strategies will be effective" (265). They explain what kinds of knowledge are useful to policy-makers, argue that such knowledge can be generated through case studies, and detail how policymakers use this knowledge as an aid to decision-making.

Some readers may find George and Bennett's focus on policy-relevant research too restrictive, and certainly in an age when Perestroika has entered political science there is room to debate models of relevance and discuss the different ways social inquiry might meaningfully inform social life. George and Bennett, by making the question of social and political relevance explicit, have opened the door to precisely such debate. Their work, taken as a whole, suggests that we do not want (as some commentators have suggested) to *replace* method-driven research with problem-driven research. Rather, we want research that is driven by important problems *and* by a sophisticated knowledge of methods. We do not

have to choose between research that is relevant and scholarship that is methodologically sound—there is no reason we cannot have both. George and Bennett’s book, in making their philosophical commitments explicit and connecting them both to methods and to questions of relevance, provide a model of scholarship that deserves to be emulated in discussions of research methodology in the social sciences.

References

Holloway, David. 1996. *Stalin and the Bomb: The Soviet Union and Atomic Energy, 1939-1956*. New Haven: Yale University Press.

Advancing the Dialogue on Qualitative Methods

Andrew Bennett

Georgetown University

bennetta@gunet.georgetown.edu

I thank my colleagues for their serious and careful reading of Alexander George’s and my book, and Jack Levy and John Gerring for organizing this symposium.¹ Publishing a book is always something of a Rohrschach test—you offer up your “ink blots” and wait to see which of the points you were less sure of or committed to will be pounced upon or embraced, which of the arguments you felt the most defensible will garner praise or come in for unexpected criticism, and what patterns will emerge. I am pleased that there seems to be considerable convergence among the critiques on important issues that our book got right, and interested to find a bit more divergence on what it could have done better or differently. Given that research methods are fraught with trade-offs, to achieve any consensus is an accomplishment, and clarifying which points lack consensus helps advance the discourse on research methods. Thus, like our critics, I will focus only briefly on the considerable areas of consensus and direct most of my attention to areas of divergence in the hope that this will best move the dialogue forward.

Points of Consensus

I am pleased to see that in our colleagues’ views we achieved several of the goals of our book, including clarifying techniques of process tracing (Dessler, Levy, Odell), highlighting the value of combining case study, statistical, and formal methods (Drezner, Levy), and elucidating the connections between case study methods and the contemporary philosophy of science (Dessler, Drezner, Levy, Odell). There also appears to be a general consensus that we have appropriately identified the comparative advantages and limitations of case study methods, although I address below several critiques that take issue with our book on some of the particulars. This rough consensus is perhaps best expressed in Odell’s summary: case studies are particularly strong at helping to develop new concepts and hypotheses, explore hypothesized

causal mechanisms, and accommodate complex forms of causal relations, and while issues of case selection, theory testing, and generalization remain challenging in qualitative work, some of the standard critiques of case study methods on these issues have missed the mark.

This is an important swath of agreement, but I would highlight as well our discussion of typological theorizing, which Odell identifies to my surprise as one of our book’s “most distinctive ideas.” This section of the book did not engender specific praise or criticism in the other critiques and has thus far received relatively little attention in print, with the important exception of an excellent article by Colin Elman (2005). While Odell is right that few have yet used this approach in print, I expect that this will change after our book has been out for a longer time, and I continue to find implicit uses of typological theorizing that could be made even stronger through more explicit and systematic use of the techniques we discussed (see, for example, Edelstein 2004).

Points of Contention

Important points of consensus notwithstanding, I focus on five critiques of our book. These include the criticisms that the research examples we used to illustrate our arguments are biased toward international security issues and micro-level process tracing; that we were too critical of statistical methods, formal modeling, and comparative methods; that we gave insufficient attention to the challenges of generalizing from and testing theories with case studies; that we should have given even more emphasis to the historical explanation of specific cases; and that laws as well as mechanisms play roles in causal explanation.

Odell and Drezner rightly note that the examples we chose to illustrate our arguments over-represented micro-level process tracing in international security research. As Odell argues, we could have included more examples from the voluminous qualitative literatures on international political economy, environmental studies, and comparative politics (and, he could have added, economics, sociology, public policy, business management, education, and many other fields). There is a method to our myopia, however. We chose examples we know well, because it is difficult to render methodological judgments on research designs without knowing a lot about the theories being examined and the empirical domain being explored.

This does not mean, as Drezner implies, that we think that process tracing is not applicable to macro-level examination or structural theories. Indeed, we explicitly stated in the context of process tracing on page 142 that “macrosocial mechanisms can be posited and tested at the macro level.” Brian Downing’s book on military revolutions and political change is one of many examples that we could (should) have cited (Downing 1992). Drezner himself gives an example in his reference to Jared Diamond’s work, and Levy approvingly notes our argument that case study methods are compatible with nearly any theoretical orientation. We added the important caveat that causal inferences based on process tracing are subject to critique if it can be shown that at a still more micro-level of analy-

sis or smaller unit of time hypothesized processes or mechanisms did not operate as claimed, but this applies to agent-based as well as structural theories. While for reasons of parsimony social scientists might simplify our models of underlying mechanisms, we should not too readily adopt “as if” assumptions that are demonstrably false.

A second line of critique is that our book is too critical of statistical, formal, and comparative methods. Having devoted a good deal of our book to addressing misunderstandings about case study methods that have arisen among scholars in other methodological traditions, we faced a difficult challenge in working to avoid mistakes in assessing alternative methods, including formal and statistical techniques, in which we are not experts. The comparative advantages of alternative methods are a central issue to methodological practice, however, and can only be discussed in the context of assessing one method relative to another. I viewed our goal on this issue as moving the discussion forward even at the risk of failing to get the comparative advantages of each method exactly right, and I look forward to a wider dialogue in which experts on statistical and formal methods detail their views on the comparative advantages and limits of alternative methods.

In this context Drezner suggests we exhibited a double standard in urging scholars to use case studies to assess complex middle-range theories, while critiquing some formal models as being “too clever by half” and positing processes more complex than those that individuals are likely to use in real world decision-making. Our point here was not that individuals are not clever or complex, or that formal models should not aspire to capture the complexities of actual decision-making, but quite the opposite. Rather, models, formal or otherwise, should not posit decision processes that are more complex than or otherwise notably different from those that individuals are likely to have used. A well-known example is that models that suggest that individuals “satisfice” are in many contexts more accurate than those that posit that individuals “maximize.” The standard is the same for formal or ordinary language theories: we should seek models whose assumptions are consistent with the processes through which individuals decide, and simplifications for the purposes of parsimony and complexifications for the sake of logical consistency should be acknowledged as such.

Drezner also argues that on page 21 of our book we wrongly critiqued statistical methods for lacking means of generating new hypotheses, and in particular, identifying new variables. Our point was that while the inductive and deductive ways of reasoning that can help identify new hypotheses and variables are open to researchers in all methodological approaches, these ways of reasoning lie outside the phase of a statistical research project in which the researcher is actually running the numbers. Regression analysis can help identify deviant cases that may be fruitful locations to look for omitted variables, but by definition it provides few clues as to why a case is an outlier. Drezner’s own example of the State Failure Task force indicates why this is so. Drezner states that the Task Force amassed more than a thousand possible independent variables and then proceeded to assess which of

them were most highly correlated with the outcome of interest. The risk of this kind of specification search, well known to practitioners of regression analysis, is that by definition one out of twenty variables (or in this example 50 or more variables) should prove “significant” at the 5% confidence level strictly by chance. I agree with Drezner that quantitative methods can and have pointed out empirical puzzles that have led to new theory-building, and we made this point ourselves in our chapter on the Democratic Peace research program. It is notable, however, that as the research program turned from testing whether democracies were unlikely to make war upon one another to developing theories on why this might be so, researchers largely shifted to studying individual cases and formal models of the deductive logic of posited mechanisms, although statistical studies continue to have an important role in testing the new theories developed from these different modes of analysis.

On a related issue, Odell rightly points out that both case study and statistical methods have advantages and disadvantages regarding different aspects of theory testing. He argues that statistical tests typically have a higher number of observations, stronger claims of representativeness, and a high degree of precision, while case studies can often consider a greater number and diversity of rival explanations. (I would add that case studies can also help assess endogeneity/feedback loops/causal direction, and potential spuriousness, and they may be able to operationalize variables in more conceptually valid ways.) Odell suggests that “on balance” scholars should avoid claiming to have tested a theory with case studies, but in my view the definitiveness of theory testing procedures depends on many factors in addition to those that Odell identifies, and varies within each methodological tradition as well as between the traditions. A theory’s failure in a most likely case, for example, is a strong finding, and its success in such a case is not very revealing.

I also take issue with Odell’s suggestion that we were too critical of comparative methods. One goal of our book was to correct for the earlier qualitative literature’s overemphasis on comparative methods relative to within-case methods of analysis. James Mahoney’s discussion of Theda Skocpol’s *States and Social Revolutions*, for example, shows that while many of the early critiques of this work focused on Skocpol’s use of Mill’s methods of comparison, Skocpol undertook considerable process tracing and was much the stronger for it (Mahoney 1999, 2000). Our point was not that comparative methods are generally weak, and we addressed in detail the uses of different kinds of case comparisons (most similar, least similar, deviant, most/least likely). Rather, our argument is that the well-known limitations of exclusive reliance on comparative methods (such as the absence of cases that are perfectly similar in all but one independent variable) can be ameliorated through the use of within-case methods of analysis like process tracing. In particular, we emphasized that typological theorizing is a means of integrating within-case and comparative methods.

I think Odell would agree with this central point, but his critique usefully draws attention to the fact that various quali-

tative methods and methodologists weigh the costs and benefits of within-case analysis and cross-case comparisons quite differently. Charles Ragin's "fuzzy set" methods, for example, put less emphasis on within-case analysis, in part because he gears these methods toward analysis of between 10 and 50 cases, too many to process trace in detail and too few for strong results from standard statistical analysis. The tradeoffs involved in this approach are quite different from those raised by our discussion of typological theorizing, and deserve more attention (Bennett and Elman 2006; Ragin 2000, 2005; Seawright 2005).

Odell argues as well that we did not adequately address the problem of drawing generalizations from case studies. We suggested that it is often useful to narrow the focus to a subclass of events rather than trying to generalize to a broad population that may exhibit heterogeneity, but Odell argues that this merely transfers the problem of representativeness down to the level of the subclass. I offer several rejoinders. Depending on how narrowly the subclass is defined and how many cases fit it, there is likely to be greater homogeneity among cases in the subclass than in the population. The number of cases in the subclass may even be sufficiently small that we can study many or all of the cases fitting this type. Thus, case studies often achieve strong and deep generalizations over narrow and precisely specified populations, while statistical studies often achieve less detailed generalizations over wider populations, and neither form of knowledge is inherently superior for all purposes. Least likely cases that fit a theory and most likely cases that do not provide grounds for generalization as well. Also, it is hard to know whether cases are representative of a population or a subclass, regardless of our methods, unless we have some understanding of the causal mechanisms involved. Darwin, as it turned out, was not just studying a few un-representative species on a quirky island, but uncovering a causal mechanism (evolutionary selection) that had broad applicability. He couldn't know of what population his species were representative until he understood the mechanism. Generalization does indeed remain a central challenge for case studies, and we might have outlined more clearly in our book this last path of generalizing through a superior understanding of the mechanisms in a case.

Conversely, Dessler worries that we may have focused too much on the goal of theoretical generalization. He emphasizes that historical explanation of individual cases is an important epistemic goal in its own right. I certainly agree that historical explanation of cases is a worthy goal in itself. Moreover, it is hard to generalize from the results of a case unless you first get its historical explanation right. Case study methods are useful for both the explanatory and generalizing goals that Odell and Dessler emphasize, although any particular research project may prove stronger at achieving one of these objectives than the other.

Finally, Dessler reminds us that both laws AND mechanisms play a role in theoretical explanation. I agree that to explain by reference to a mechanism is in many respects the same as explanation by reference to a law. The mechanism ex-

planation is perhaps narrower and more contingent, but in form it is not different from explanation via reference to a law: both statements argue that in a specified context (one narrower than the other) the mechanism/law renders the outcome as something that was to be expected. The subtle difference between the two forms of explanations is that most discussions of explanation via laws more readily admit "as if" assumptions about processes or mechanisms, whereas mechanism-based explanations admit to being subject to refutation if it can be demonstrated, often at a lower level of analysis or finer degree of detail, that the posited mechanism was not in operation.

This issue parallels Levy's argument that statistical tests can also get at causal mechanisms because the posited mechanisms should have observable implications for statistical distributions. Indeed, they should, and statistical tests can in this way bolster our confidence in theorized mechanisms, and their generalizability, but correlational findings also have a greater "as if" quality than do explanations based on process tracing within a case. This is not to say that the latter method is free from risks of inferential error, just that it is a different form of inference. As Dessler notes, identifying the underlying mechanism (provided we do it correctly, which cannot be known with certainty) repairs the deficiency of the Deductive-Nomological model by showing that the law provides a causal explanation rather than an accidental association.

Next Steps in the Qualitative Methods Dialogue

Odell is certainly right that neither our book nor any other single text provides a stand-alone basis for a graduate course in contemporary qualitative methods. We addressed important issues, like interview techniques, only in passing, and many important developments in qualitative methods were published too recently for us to fully incorporate them into our book (for a current syllabus on the subject, see <http://www.asu.edu/clas/polisci/cqrm>). These include books by Brady and Collier (2004), Goertz (2006), and Gerring (2007), as well as many articles by these and other authors. Other important innovations are ongoing, including analyses of how to carry out multimethod work that combines case studies with statistical analysis and/or formal models (Lieberman 2004, Gerring and Seawright forthcoming, Bennett 2002), clarification of the Bayesian logic that informs process tracing practices (Bennett forthcoming 2006), examination of the role of qualitative methods in addressing causal complexity (Bennett and Elman, forthcoming), and an effort by Andrew Moravcsik of Princeton to clarify procedures for weighing the value and improving the web accessibility of archival materials (a project much like that which Drezner urges herein). Other topics raised by Drezner (agent-based modeling, cognitive neuroscience, archaeological techniques, and online simulation) and Levy (combinations of experimental and qualitative techniques) deserve attention as well.

This unusually torrid pace of innovation is at one and the same time developing a consensus on qualitative methods and pushing the envelope on new techniques. Just as our book builds on decades of practice with the methods devel-

oped and refined by George, Lijphart, Eckstein, and others, we will no doubt learn more about the comparative advantages of recent methodological innovations from attempts to put them into practice. In the not-too-distant future these efforts may coalesce into more standardized and comprehensive (though still evolving) textbooks in qualitative methods analogous to those found in graduate courses in statistics. Brady and Collier's call for "diverse tools and shared standards" is already a far more substantial reality than when we began our book.

Notes

¹ Alexander George decided not to co-author this response, as he was unable to attend the APSA roundtable that originated it and has had some health problems as he approaches his 86th birthday. He and his wife Juliette are also busily preparing to move to the Seattle area, where their daughter lives, but I expect they would welcome hearing from old friends. I have occasionally used the pronouns "we" or "our" in this article to refer to our co-authored book or the thinking behind it, but I remain solely responsible for the content of this response.

References

- Bennett, Andrew. 2002. "Where the Model Frequently meets the Road: Combining Statistical, Formal, and Case Study Methods." Presented at the American Political Science Association conference in Boston; Currently under revision with Bear Braumoeller as co-author.
- Bennett, Andrew. 2006. "Stirring the Frequentist Pot with a Dash of Bayes." Forthcoming in *Political Analysis*.
- Bennett, Andrew and Colin Elman. 2006. "Qualitative Research: Recent Developments in Case Study Methods." *Annual Review of Political Science* 9, 455-76.
- Bennett, Andrew and Colin Elman. 2006. "Complex Causal Relations and Case Study Methods: The Example of Path Dependence." Forthcoming in *Political Analysis*.
- Brady, Henry E. and David Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman and Littlefield.
- Downing, Brian. 1992. *The Military Revolution and Political Change*. Princeton NJ: Princeton University Press
- Eckstein, Harry. 1975. "Case Study and Theory in Political Science." In *Handbook of Political Science* 7, eds. Fred I. Greenstein and Nelson W. Polsby, (Reading, MA: Addison-Wesley), 79-137.
- Edelstein, David M. 2004. "Occupational Hazards: Why Military Occupations Succeed or Fail." *International Security* 29:1, 49-91.
- Elman, Colin. 2005. "Explanatory Typologies in Qualitative Studies of International Politics." *International Organization* 59:2, 293-326.
- George, Alexander L. 1979. "Case Studies and Theory Development: The Method of Structured, Focused Comparison. In *Diplomacy: New Approaches in History, Theory and Policy*, ed. G Lauren, (New York: Free Press), 43-68.
- Gerring, John. 2007. *Case Study Research: Principles and Practices*. Cambridge: Cambridge University Press.
- Gerring, John and Jason Seawright. Forthcoming. "Selecting Cases in Case Study Research: A Menu of Options."
- Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press.
- Lieberman, Evan. 2004. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review*

99:3, 435-52.

- Mahoney, James. 1999. "Nominal, Ordinal and Narrative Appraisal in Macro-Causal Analysis: The Case of Theda Skocpol." *American Journal of Sociology* 104:4, 1154-96.
- Mahoney, James. 2000. "Strategies of Causal Inference in Small-N Analysis." *Sociological Methods and Research* 28:4, 387-424.
- Ragin, Charles. 2000. *Fuzzy Set Social Science*. Chicago: Chicago University Press.
- Ragin, Charles. 2005. "Core versus Tangential Assumptions in Comparative Research." *Studies in Comparative International Development* 40:1, 33-38.
- Seawright, Jason. 2005. "Qualitative Comparative Analysis vis-à-vis Regression." *Studies in Comparative International Development* 40:1, 3-26.

Book Notes

Book descriptions are excerpted from publisher's web sites. If you would like to recommend a book to be included in this section, email Joshua C. Yesnowitz, the assistant editor of QM, at jcyesnow@bu.edu.

- Biddle, Stephen D. 2004. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton: Princeton University Press. \$37.50 hardcover.

In war, do mass and materiel matter most? Will states with the largest, best equipped, information-technology-rich militaries invariably win? The prevailing answer today among both scholars and policy-makers is yes. But this is to overlook force employment, or the doctrine and tactics by which materiel is actually used. In a landmark reconception of battle and war, this book provides a systematic account of how force employment interacts with materiel to produce real combat outcomes. Stephen Biddle argues that force employment is central to modern war, becoming increasingly important since 1900 as the key to surviving ever more lethal weaponry. Technological change produces opposite effects depending on how forces are employed; to focus only on materiel is thus to risk major errors with serious consequences for both policy and scholarship. In clear, fluent prose, Biddle provides a systematic account of force employment's role and shows how this account holds up under rigorous, multimethod testing. The results challenge a wide variety of standard views, from current expectations for a revolution in military affairs to mainstream scholarship in international relations and orthodox interpretations of modern military history.

- Chernoff, Fred. 2005. *The Power of International Theory: Reforging the Link to Foreign Policy-Making through Scientific Enquiry*. London: Routledge. \$115.00 hardcover.

The discipline of International Relations was created with a purpose of helping policy-makers to build a more peaceful and just world. However, many of the current trends, post-positivism, constructivism, reflectivism, and post-modernism share a conception of international theory that is inherently incapable of offering significant guidance to policy-makers. *The Power of International Theory* critically examines these approaches and offers a novel conventional-causal alternative that allows the re-forging of a link between IR theory and policy-making. While recognizing the criticisms of earlier forms of positivism and behavioralism, the book defends holistic testing of empirical principles, methodological pluralism, criteria for choosing the best theory, a notion of "causality," and a limited form of prediction, all of which are needed to guide policy-makers. This

book will be an invaluable text for advanced students and researchers in the fields of international relations theory and the philosophy of social science.

Eden, Lynn. 2003. *Whole World on Fire: Organizations, Knowledge, and Nuclear Weapons Devastation*. Ithaca: Cornell University Press. \$24.95 paperback.

Whole World on Fire focuses on a technical riddle wrapped in an organizational mystery: How and why, for more than half a century, did the U.S. government fail to predict nuclear fire damage as it drew up plans to fight strategic nuclear war? U.S. bombing in World War II caused massive fire damage to Hiroshima and Nagasaki, but later war plans took account only of damage from blast; they completely ignored damage from atomic firestorms. Recently a small group of researchers has shown that for modern nuclear weapons the destructiveness and lethality of nuclear mass fire often—and predictably—greatly exceeds that of nuclear blast. This has major implications for defense policy: the U.S. government has underestimated the damage caused by nuclear weapons, Lynn Eden finds, and built far more warheads, and far more destructive warheads, than it needed for the Pentagon's war-planning purposes. How could this have happened? The answer lies in how organizations frame the problems they try to solve.

In a narrative grounded in organization theory, science and technology studies, and primary historical sources (including declassified documents and interviews), Eden explains how the U.S. Air Force's doctrine of precision bombing led to the development of very good predictions of nuclear blast—a significant achievement—but for many years to no development of organizational knowledge about nuclear fire. Expert communities outside the military reinforced this disparity in organizational capability to predict blast damage but not fire damage. Yet some innovation occurred, and predictions of fire damage were nearly incorporated into nuclear war planning in the early 1990s. The author explains how such a dramatic change almost happened, and why it did not. *Whole World on Fire* shows how well-funded and highly professional organizations, by focusing on what they do well and systematically excluding what they don't do well, may build a poor representation of the world—a self-reinforcing fallacy that can have serious consequences. In a sweeping conclusion, Eden shows the implications of the analysis for understanding such things as the sinking of the Titanic, the collapse of the Tacoma Narrows Bridge, and the poor fireproofing in the World Trade Center.

Fullbrook, Edward (ed). 2003. *The Crisis in Economics: The Post-autistic Economics Movement: The First 600 Days*. London: Routledge. \$41.95 paperback.

Economics can be pretty boring. Drier than Death Valley, the discipline is obsessed with mathematics and compounds this by arrogantly assuming its techniques can be brought to bear on the other social sciences. It wasn't going to be long, therefore, before students started complaining. The vast majority have voted with their feet and signed up for business and management degrees, but in the past two years there has grown an important new movement that has decided to tackle those who think they run economics head-on. This is the Post-autistic Economics Network. The PAE Network started in France and has spread first to Cambridge and then other parts of the world. The name derives from the fact that mainstream economics has been accused of institutional autism, i.e. qualitative impairment of social interaction, failure to develop peer relationships and lack of emotional and social reciprocity. In short, economics has lost touch with reality and has become way too abstract. This book charts the impact the PAE Network has had so far and constitutes a

manifesto for a different kind of economics—it features key contributions from all the major voices in heterodox economics including Tony Lawson, Deirdre McCloskey, Geoff Hodgson, Sheila Dow and Warren Samuels.

Gallagher, Nancy W. 2003. *The Politics of Verification*. Baltimore: Johns Hopkins University Press. \$25.00 paperback.

How to evaluate compliance is among the most difficult questions that arise during treaty negotiations and ratification debates. Arguments over verification principles and procedures are increasingly common for accords about the environment, human rights, and economics, but they have been especially important in the arena of national security. Nancy Gallagher explains, "In a world in which states face conflicting pressures to maximize military capabilities and negotiate mutual restraints, the prospects for arms control often hinge on verification... In the standard American formulation, verification is the 'critical element of arms control.'"

In *The Politics of Verification*, Gallagher explores the causes of verification controversies and the processes through which they are perpetuated or provisionally resolved. By examining nuclear test ban negotiations from the Eisenhower through the Clinton administrations, Gallagher finds that the assumptions about verification that have dominated U.S. policy shape domestic debates in ways that hinder stable agreement on significant test restrictions. She focuses on the dynamic interconnections between domestic and international politics, and analyzes the slow process of coalition building when conflicting interests and ideas create divisions both among and inside states. Gallagher concludes that the end of the Cold War has altered the arms control context without resolving basic questions about the appropriate amount and type of verification. Thus, the negotiation and ratification of major cooperative accords will continue to be shaped by verification compromises and coalitions.

Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press. \$27.95 paperback.

Concepts lie at the core of social science theory and methodology. They provide substance to theories; they form the basis of measurement; they influence the selection of cases. *Social Science Concepts: A Users Guide* explores alternative means of concept construction and their impact on the role of concepts in measurement, case selection, and theories. While there exists a plethora of books on measurement, scaling, and the like, there are virtually no books devoted to the construction and analysis of concepts and their role in the research enterprise. *Social Science Concepts: A Users Guide* provides detailed and practical advice on the construction and use of social science concepts; a Web site provides classroom exercises. It uses a wide range of examples from political science and sociology such as revolution, welfare state, international disputes and war, and democracy to illustrate the theoretical and practical issues of concept construction and use. It explores the means of constructing complex, multi-level, and multidimensional concepts. In particular, it examines the classic necessary and sufficient condition approach to concept building and contrasts it with the family resemblance approach. The consequences of valid concept construction are explored in both qualitative and quantitative analyses. *Social Science Concepts: A Users Guide* will prove an indispensable guide for graduate students and scholars in the social sciences. More broadly, it will appeal to scholars in any field who wish to think more carefully about the concepts used to create theories and research designs.

Lischer, Sarah. 2005. *Dangerous Sanctuaries: Refugee Camps, Civil War, and the Dilemmas of Humanitarian Aid*. Ithaca, NY: Cornell University Press. \$35.00 cloth.

Since the early 1990s, refugee crises in the Balkans, Central Africa, the Middle East, and West Africa have led to the international spread of civil war. In Central Africa alone, more than three million people have died in wars fueled, at least in part, by internationally supported refugee populations. The recurring pattern of violent refugee crises prompts the following questions: Under what conditions do refugee crises lead to the spread of civil war across borders? How can refugee relief organizations respond when militants use humanitarian assistance as a tool of war? What government actions can prevent or reduce conflict?

To understand the role of refugees in the spread of conflict, Sarah Kenyon Lischer systematically compares violent and nonviolent crises involving Afghan, Bosnian, and Rwandan refugees. Lischer argues against the conventional socioeconomic explanations for refugee-related violence—abysmal living conditions, proximity to the homeland, and the presence of large numbers of bored young men. Lischer instead focuses on the often-ignored political context of the refugee crisis. She suggests that three factors are crucial: the level of the refugees' political cohesion before exile, the ability and willingness of the host state to prevent military activity, and the contribution, by aid agencies and outside parties, of resources that exacerbate conflict. Lischer's political explanation leads to policy prescriptions that are sure to be controversial: using private security forces in refugee camps or closing certain camps altogether. With no end in sight to the brutal wars that create refugee crises, *Dangerous Sanctuaries* is vital reading for anyone concerned with how refugee flows affect the dynamics of conflicts around the world.

Monroe, Kristen R. (ed). 2005. *Perestroika!: The Raucous Rebellion in Political Science*. New Haven: Yale University Press. \$35.00 paperback.

This superb volume describes the events and ramifications of a revolt within the political science discipline that began in 2000 with a disgruntled e-mail message signed by one "Mr. Perestroika." The message went to seventeen recipients who quickly forwarded it to others, and soon the Perestroika revolt became a major movement calling for change in the American political science community. What is the Perestroika movement? Why did it occur? What has it accomplished? What remains to be done? Most important, what does it tell us about the nature of political science, about methodological pluralism and diversity, about the process of publishing scholarly work, and about graduate education in the field? The contributors to the book—thoughtful political scientists who offer a variety of perspectives—set the Perestroika movement in historical and comparative contexts. They address many topics related to heart of the debate—a desire for tolerance of methodological diversity—and assess the changes that have come in the wake of Perestroika. For political scientists and their graduate students, and for those interested in the history or sociology of social sciences, this volume is essential reading.

Rihoux, Benoit and Heike Grimm (eds). 2005. *Innovative Comparative Methods for Policy Analysis: Beyond the Qualitative-Quantitative Divide*. Berlin: Springer. \$89.95 hardcover.

Innovative Comparative Methods for Policy Analysis aims to provide a decisive push to the further development and application of innovative and specific comparative methods for the improvement of policy analysis. To take on this challenge, this volume brings together methodologists and specialists from a broad range of social

scientific disciplines and policy fields. The work further develops methods for systematic comparative cases analysis in a small-N research design, with a key emphasis laid on policy-oriented applications. *Innovative Comparative Methods for Policy Analysis* is clearly both a social scientific and policy-driven endeavor; on the one hand, the book engages in an effort to further improve social scientific methods, but on the other hand this effort also intends to provide useful, applied tools for policy analysts and the "policy community" alike. Though quite a variety of methods and techniques are touched upon in this volume, its focus is mainly laid on two recently developed research methods/techniques which enable researchers to systematically compare a limited number of cases; Qualitative Comparative Analysis (QCA) and Fuzzy-Sets (FS).

Stinchcombe, Arthur M. 2005. *The Logic of Social Research*. Chicago: University of Chicago Press. \$20.00 paperback.

Arthur L. Stinchcombe has earned a reputation as a leading practitioner of methodology in sociology and related disciplines. Throughout his distinguished career he has championed the idea that to be an effective sociologist, one must use many methods. This incisive work introduces students to the logic of those methods. *The Logic of Social Research* orients students to a set of logical problems that all methods must address to study social causation. Almost all sociological theory asserts that some social conditions produce other social conditions, but the theoretical links between causes and effects are not easily supported by observation. Observations cannot directly show causation, but they can reject or support causal theories with different degrees of credibility. As a result, sociologists have created four main types of methods that Stinchcombe terms quantitative, historical, ethnographic, and experimental to support their theories. Each method has value, and each has its uses for different research purposes. Accessible and astute, *The Logic of Social Research* offers an image of what sociology is, what it's all about, and what the craft of the sociologist consists of.

Tetlock, Philip. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton: Princeton University Press. \$35.00 hardcover.

The intelligence failures surrounding the invasion of Iraq dramatically illustrate the necessity of developing standards for evaluating expert opinion. This book fills that need. Here, Philip E. Tetlock explores what constitutes good judgment in predicting future events, and looks at why experts are often wrong in their forecasts. Tetlock first discusses arguments about whether the world is too complex for people to find the tools to understand political phenomena, let alone predict the future. He evaluates predictions from experts in different fields, comparing them to predictions by well-informed laity or those based on simple extrapolation from current trends. He goes on to analyze which styles of thinking are more successful in forecasting. Classifying thinking styles using Isaiah Berlin's prototypes of the fox and the hedgehog, Tetlock contends that the fox—the thinker who knows many little things, draws from an eclectic array of traditions, and is better able to improvise in response to changing events—is more successful in predicting the future than the hedgehog, who knows one big thing, toils devotedly within one tradition, and imposes formulaic solutions on ill-defined problems. He notes a perversely inverse relationship between the best scientific indicators of good judgement and the qualities that the media most prizes in pundits—the single-minded determination required to prevail in ideological combat. Clearly written and impeccably researched, the book fills a huge void in the literature on evaluating expert opinion.

Topper, Keith. 2005. *The Disorder of Political Inquiry*. Cambridge: Harvard University Press. \$45.00 hardcover.

In the past several years two academic controversies have migrated from the classrooms and courtyards of college and university campuses to the front pages of national and international newspapers: Alan Sokal's hoax, published in the journal *Social Text*, and the self-named movement, "Perestroika," that recently emerged within the discipline of political science. Representing radically different analytical perspectives, these two incidents provoked wide controversy precisely because they brought into sharp relief a public crisis in the social sciences today, one that raises troubling questions about the relationship between science and political knowledge, and about the nature of objectivity, truth, and meaningful inquiry in the social sciences. In this provocative and timely book, Keith Topper investigates the key questions raised by these and other interventions in the "social science wars" and offers unique solutions to them. Engaging the work of thinkers such as Richard Rorty, Charles Taylor, Pierre Bourdieu, Roy Bhaskar, and Hannah Arendt, as well as recent literature in political science and the history and philosophy of science, Topper proposes a pluralist, normative, and broadly pragmatist conception of political inquiry, one that is analytically rigorous yet alive to the notorious vagaries, idiosyncrasies, and messy uncertainties of political life.

Announcements

APSA Poster Sessions Organized by Division 46: Qualitative Methods August 31-September 3, 2006, Philadelphia, PA (Provisional pending final acceptance by authors)

- Amel F. Ahmed, University of Pennsylvania: "Comparative Historical Analysis and the Politics of Institutional Choice: Explaining Voting System Reform in 19th Century Democratizers."
- Alejandra Betanzo de la Rosa: "Minding the Gap: Connecting Theory and Data in the Field of Intergovernmental Relations."
- Olga Bogatyrenko, University of California-Davis: "Vulnerabilities of the Powerful: Using Historical Perspective to Understanding Power."
- Morris D Bidjerano, State University New York at Albany: "Who Gets What Public Policy, When and How: Reconceptualizing Power in Public Policy Making."
- Chien-peng Chung, Lingnan University, Hong Kong: "An Analytical Framework of Separatism in Modern China."
- Hristina Nikolaeva Dobрева, Simon Fraser University, Canada: "'Second Image Reversed' Reexamined: Methodological Reconsideration of Bias in Small-N Research."
- Fred Eidlin, University of Guelph, Canada: "Ideal Types and the Problem of Reification."
- Michael Javen Fortner, Harvard University: "The Mismeasure of Identity: Aligning Ontology and Methodology in Race Politics."
- Els de Graauw, University of California-Berkeley: "Conceptualizing and Typologizing Immigrant Nonprofits as Actors in American Urban Politics."
- Petra Hejnova, Syracuse University: "Women's Activism in the New Democracies: Uncovering Effects of Communist Policies on Czech Women."
- Adrian J. Lottie, Eastern Michigan University: "Racism as Epistemology."
- Margitta Maetzke, University of Bremen: "Germany, Power and

- Preferences in the Development of Welfare States: Country-Level Generalizations and its Alternatives."
- Robert Mickey, University of Michigan: "Duration, Tempo, and the Study of Macro-Political Processes."
- Christopher Newman, Roosevelt University: "Fuzzy Set Qualitative Comparative Analysis of Rebellions and Revolutions."
- Simeon C. Nichter, University of California-Berkeley: "Shaping Opportunities for Collective Action: The Case of Land Reform in Brazil."
- Hironori Sasada, University of Washington: "Ideas, Individuals, and Institutions in State Development: The Origin of the East Asian Developmental State System."
- Andrew J. Seligsohn, Hartwick College: "Values in American Politics."
- Paul Steinberg, Harvey Mudd College: "Causal Analysis in Small-N Policy Studies."
- Nicholas Toloudis, Columbia University: "Centralization Reconsidered: Moving Beyond 'Dynamic Statism.'"
- Andreas Umland: "'Conservative Revolution': Proper Name or Generic Concept?"

APSA Panels/Roundtables Created (or Co-Organized) by Division 46: Qualitative Methods August 31-September 3, 2006, Philadelphia, PA (Provisional pending final acceptance by authors)

Theoretical Synthesis: Empirical Advances in International Politics

- Chair: Peter J. Katzenstein, Cornell University
- Participants:
- Rachel Epstein, University of Denver: "The Social Context in Conditionality: Internationalizing Finance and Defense."
- Aneta Borislavova Spendzharova, University of North Carolina: "For the Market, or for 'Our Friends'? Reforming Banking Laws in Hungary and Bulgaria after 1989."
- Liesbet Hooghe, University of North Carolina: "Efficiency, Distribution, Trust and International Regime Design."
- Judith Kelley, Duke University: "Norms, Systemic Change and Instrumentalism: Explaining the Rise of Election Monitoring."
- Discussants: Jeffrey T. Checkel, University of Oslo; Milada Anna Vachudova, University of North Carolina

What has Comparative Politics Accomplished?: A Conversation Among Leading Scholars

- Chair: Richard Snyder, Brown University
- Participants: Robert A. Dahl, Yale University; David D. Laitin, Stanford University; Theda Skocpol, Harvard University; Alfred C. Stepan, Columbia University
- Discussant: Gerardo L. Munck, University of Southern California

Temporal Dimensions of Policies and Politics

- Chair: Kathleen Thelen, Northwestern University
- Participants:
- Anna M. Grzymala-Busse, University of Michigan: "Disaggregating Temporal Effects and their Policy Impact."
- Maria Victoria Murillo, Columbia University: "Winners Take All or Policy Feedback Effects?"
- Alan M. Jacobs, University of British Columbia: "When Can Governments Invest?: Institutions, Government Capacities, and Policies for the Long Term."
- Tulia G. Falleti, University of Pennsylvania: "Policies in Time: The

Long- and Short-Durée Causes of Policy-Making Processes.”
Shannon O’Neil, Harvard University: “Un-privatizing Pensions?:
The Divergent Political Effects of Social Security Reform in the
Short, Medium and Long Term.”
Discussants: Evelyne Huber, University of North Carolina; Stephen
E. Hanson, University of Washington

Descriptive Inference: Is There a Method to Describing?

Chair: Bear F. Braumoeller, Harvard University
Participants:
John Gerring, Boston University: “What is the Difference between
Good Description and Bad Description?”
Kevin M. Quinn, Harvard University: “Descriptive Statistical In-
ference.”
Jake Bowers, University of Michigan: “The Shape of Political Par-
ticipation: Descriptive Inference with Normative and Scientific
Implications.”
Discussants: Bear F. Braumoeller, Harvard University; Alexander
Wendt, Ohio State University

Multimethod Research: Best Practices, Challenges, and Exemplars

Chair: Andrew Bennett, Georgetown University
Participants: Margaret Levi, University of Washington-Seattle; Rob-
ert H. Bates, Harvard University; Hein Erich Goemans, Univer-
sity of Rochester; Kenneth A. Schultz, Stanford University; Evan
S. Lieberman, Princeton University
Discussant: Andrew Bennett, Georgetown University

Beyond Continuity: Authors Meet Critics

Chair: Kathleen Thelen, Northwestern University
Participants: Theda Skocpol, Harvard University; Isabela Mares,
Stanford University; Ellen M. Immergut, Humboldt University,
Berlin; Pauline Jones-Luong, Brown University
Discussant: Jacob S. Hacker, Yale University

Path Dependence Meets Punctuated Equilibrium: Similarities and Differences

Chair: Gary Goertz, University of Arizona
Participants: Jack A. Goldstone, George Mason University; Bryan
D. Jones, University of Washington; Paul F. Diehl, University of
Illinois, Urbana-Champaign; Sheri Berman, Barnard College; Frank
R. Baumgartner, Pennsylvania State University

Concept Analysis: Unpacking and Reconceptualizing, Clientelism, Governance, and Neoliberalism

Chair: Marcus J. Kurtz, Ohio State University
Participants:
Taylor C. Boas (and Jordan Luc Gans-Morse), University of
California-Berkeley: “From Rallying Cry to Whipping Boy: The
Concept of Neoliberalism in the Study of Development.”
Sebastian R. Karcher, Northwestern University: “Reconciling Con-
tinuous Measurement and Diminished Subtypes—The Case of the
Informal Economy.”
Aaron Schneider, IDS at the University of Sussex: “Rethinking Gov-
ernance: Shared Standards among Quantitative, Qualitative, and
Interpretive Tools.”
Nikolaos Bizziouras, Harvard University: “Reconceptualizing
Clientelism: Combining Demand- and Supply-Driven Ap-
proaches.”

Jonathan Hopkin, London School of Economics: “Conceptualizing
Clientelism: Political Exchange and Democratic Theory.”
Discussant: Marcus J. Kurtz, Ohio State University

Roundtable on the Work and Legacy of Giovanni Sartori

Chair: David Collier, University of California-Berkeley
Participants: Gary Goertz, University of Arizona; John Gerring,
Boston University; Cindy L. Skach, Harvard University; Marcus
J. Kurtz, Ohio State University
Discussants: Giovanni Sartori, Columbia University; David Collier,
University of California-Berkeley

Issues in the Ethics and Process of Data Collection and Analysis

Chair: Benjamin L. Read, University of Iowa
Participants:
Jonathan B. Isacoff, Gonzaga University: “Pragmatism, Politics,
Political Science.”
T. Allen Lambert and Enoch A. Lambert, SUNY Albany: “Method-
ological Issues in the Study of Governance: Practitioner-Oriented
Design, Qualitative Information, and User Confidence—Episte-
mological and Pragmatic Issues in Applied Research.”
Lauren M. Morris MacLean, Indiana University: “The Power of
Position: The Consideration of Power in the Process of Political
Science Research.”
Bernd Reiter, University of South Florida: “When to Stop Inter-
viewing: Applying Insights from Gadamer’s Hermeneutic Circle.”
Discussant: Benjamin L. Read, University of Iowa

Process Tracing and Causal Analysis in Small-N Research

Chair: Colin Elman, Arizona State University
Participants:
Abhishek Chatterjee, University of Virginia: “Toward An Ontology
of Case Studies (Or Why Most Existing Defenses of Qualitative
Case Studies Fail).”
Ingo Rohlfing, International University, Bremen: “Does Process
Tracing Solve the n=1 Problem of Case Study Research?”
James Mahoney, Northwestern University: “Scope: Causal and
Conceptual Homogeneity in Qualitative Research.”
Hillel David Soifer, Harvard University: “The Implications of Pro-
cess-Tracing for Case Selection in Comparative Research De-
sign.”
Oisín Tansey, Nuffield College: “Process Tracing and Elite Inter-
viewing: A Case for Nonprobability Sampling.”
Discussant: Colin Elman, Arizona State University

Multimethod Research: Principles and Prescriptions

Chair: Evan S. Lieberman, Princeton University
Participants:
Andrew Bennett, Georgetown University: “Where the Model Fre-
quently Meets the Road: Combining Statistical, Formal, and Case
Study Methods.”
Mona Lena Krook, Washington University in St. Louis: “Temporal-
ity and Causal Configurations: Combining Sequence Analysis and
Fuzzy Set/Qualitative Comparative Analysis.”
Patrick Thaddeus Jackson, American University: “A Statistician
Strikes Out: In Defense of Genuine Methodological Diversity.”
Giovanni Capocchia (and Michael Freedon), University of Oxford:
“Mixed Method Research in Comparative Politics: A Discussion
of Basic Assumptions.”
Nicholas C Wheeler, University of Virginia: “Risk vs. Uncertainty

in Political Research: Assessing the Need for Methodological Pluralism.”

Discussant: Evan S. Lieberman, Princeton University

Making Black Politics Visible Through Qualitative Methods

Chair: Dianne M. Pinderhughes, University of Illinois, Urbana-Champaign

Participants:

Lorrie A. Frasure, Cornell University: “Immigrants, Ethnic Minorities and the Logic of Suburban Institutional Interdependency.”

Megan Francis, Princeton University: “Institutionalized Injustice: The Relationship between Politics and Black Incarceration.”

Alvin B. Tillery Jr., University of Notre Dame: “The Paradox of Power: The Congressional Black Caucus and U.S. Foreign Policy Toward Africa Since the End of Apartheid.”

Dorian T. Warren, University of Chicago: “When Wal-Mart Comes to Town: Inequality, Power and the Neoliberalization of Black Politics.”

Discussants: Linda Faye Williams, University of Maryland; Dianne M. Pinderhughes, University of Illinois, Urbana-Champaign

The Politics of Ideas: Theory, Method, and Analysis

Chair: Robert H. Cox, University of Oklahoma

Participants:

Vivien A. Schmidt, Boston University: “Give Peace a Chance: Reconciling Four (not Three) ‘New Institutionalisms.’”

Colin Hay, University of Birmingham: “Constructivist Institutionalism—Or, Why Interests into Ideas Don’t Go.”

Craig Parsons, University of Oregon: “Modernity, Ideas, and Supra-nationality.”

Andrew Rich, City College of New York & CUNY Grad Center: “Ideas versus Expertise: Think Tanks and the Organizations of Information in American Policymaking.”

Daniel Wincott, University of Birmingham: “New Social Risks and the Changing Welfare State: Ideas, Policy Drift and Early Childhood Education and Care Policies in Australia, the UK and the USA.”

Discussants: Daniel Beland, University of Calgary

Roundtable on Gary Goertz, “Social Science Concepts: A User’s Guide”

Chair: Melani Cammett, Brown University

Participants: Steven Levitsky, Harvard University; David Collier, University of California-Berkeley; Theda Skocpol, Harvard University; Gary Goertz, University of Arizona

Rethinking Interpretation and the Interpreter

Chair: Audie Klotz, Syracuse University

Participants:

Matthew Hoffman, University of Delaware: “Modeling Constructivist Insights? Agent-Based Modeling and Social Dynamics.”

J. Samuel Barkin, University of Florida: “What Defines Research as Qualitative?”

Anna Leander, Copenhagen Business School: “The ‘Real Politics of Reason’: Thinking International Relations through Bourdieu’s Fields, Habitus and Practice.”

Kevin Dunn, Hobert and William Smith College: “Examining Historical Representations.”

Patricia Goff, Wilfred Laurier University, Canada: “Revisiting the Question of Rigor in Qualitative Methods.”

Discussants: Ted Hopf, Ohio State University; Nina Tannenwald, Brown University

Social Inquiry and Political Knowledge

Chair: Martin O. Heisler, University of Maryland

Participants: Macartan Humphreys, Columbia University; John Gerring, Boston University; Stefano Guzzini, Uppsala University/Danish Institute for International Studies; Martin O. Heisler, University of Maryland

(De)coding Discourse

Chair: Mark Bevir, University of California-Berkeley

Participants:

Emily D. Shaw, University of California-Berkeley: “(De)coding Issue Frames: Emergent Code Identification in Content Analysis.”

Elisabeth Anker (and Matthew Scherer), University of California-Berkeley: “Political Theory/Political Discourse: Constructing Moral and Religious Legitimacy for Public Action.”

Bethany Albertson, University of Chicago: “Coded Communication and Religious Appeals in American Politics.”

Kevin Jay Wallsten, University of California-Berkeley: “What Are the Blogs Telling Us? Exploiting Blogs as a Source of Data.”

Discussants: Laura Stoker, University of California-Berkeley; Mark Bevir, University of California-Berkeley

The Making of Modern Political Science

Chair: Shannon C. Stimson, University of California-Berkeley

Participants:

James Farr, University of Minnesota: “The Historical Science(s) of Politics: The Principles, Association, and Fate of an American Discipline.”

John G. Gunnell, State University New York at Albany

Robert Kaufman Adcock, Stanford University

Mark Bevir, University of California-Berkeley

Examples of Multi-Method Research

Chair: Rose McDermott, University of California-Santa Barbara

Participants:

Daniel P. Carpenter, Harvard University: “The Calculus of Patient Advocacy in FDA Drug Review: Stories and Models.”

Dara Z. Strolovitch, University of Minnesota-Twin Cities: “Using Multiple Methods to Study Multiple Marginalization: Qualitative, Quantitative, and Normative Approaches to Studying the Intersectional Politics of Racial, Gender, and Income Inequalities.”

Eric Brahm, University of Colorado: “A Multi-Method Exploration of Truth Commissions.”

Laura Sjoberg (and Christopher Marcoux), Harvard University: “Was Iraq a Threat to International Peace and Security? Methodological Insights to Explore the Question.”

Nathaniel Beck, New York University: “Qualitative and Quantitative Methods: Can They Be Joined?”

Discussant: Rose McDermott, University of California-Santa Barbara

Ethnographic Methods and Knowledge about Politics

Chair: Jessica Allina-Pisano, Colgate University

Participants:

Lorraine Bayard de Volo, University of Kansas: “Ethnography, Politics, and Power Relations: Participant Observations with Nicara-

guan Mothers and U.S. Casino Waitresses.”
Tim Pachirat, Yale University: “Ethnography from Below? Reflections from an Industrialized Slaughterhouse on Perspective, Power, and the Ethnographic Voice.”
Edward Schatz, University of Toronto: “The Problem with the Toolbox Metaphor: Ethnography and the Limits to Multiple Methods Research.”
Discussants: James C. Scott, Yale University; Dvora Yanow, Vrije Universiteit, Amsterdam

Coding International Treaties

Chair: Ronald B. Mitchell, University of Oregon
Participants:
Barbara Koremenos, University of Michigan: “The Challenge of Characterizing the Cooperation Problem.”
Aslaug Asgeirs-dottir, Bates College: “Coding Maritime Boundary Agreements: The 1982 United Nations Law of the Sea and International Cooperation around Common Ocean Areas.”
Brett Ashley Leeds, Rice University: “The Challenges of Collecting and Coding Primary Documents: Lessons from the ATOP Project.”
Monty G. Marshall, George Mason University: “Objectifying the Subjective: Confessions of a Data Junkie.”
Steven B Rothman (and Ronald B. Mitchell), University of Oregon: “Creating Large-N Databases from Qualitative Information: Lessons from International Environmental Agreements.”
Discussant: Paul F. Diehl, Univ. of Illinois, Urbana-Champaign

Interpretive Methods in Practice

Chair: Lee Ann Fujii, George Washington University
Participants:
Eric M. Blanchard, University of Southern California: “Beyond Geertz: Discourse Analysis as Qualitative Method.”
Lee Ann Fujii (and Kara Heitz), George Washington University: “Using Personal Narratives as Evidence: Insights from Fieldwork in Rwanda.”
Jurgen Petersen, Hauptseite Universität Zürich: “Studying Actors’ Cultural Theories: Approaches and Practice.”
Dvora Yanow, Vrije Universiteit-Amsterdam: “Improv and Interpretive methods.”
Discussant: Cecelia Lynch, University of California-Irvine

Norms vs. Ideas: What’s at Stake for Comparative Politics and International Relations

Chair: Jeffrey T. Checkel, University of Oslo
Participants: Regina M. Abrami, Harvard University; Sheri Berman, Barnard College; David Woodruff, Harvard University; Ted Hopf, Ohio State University; Yoshiko M. Herrera, Harvard University

The Methods Café at APSA: Consult a Specialist

Date: TBA (See <http://www.apsanet.org> as conference approaches)

Organizers: Peregrine Schwartz-Shea, University of Utah; Dvora Yanow, Vrije Universiteit, Amsterdam.

At the Methods Café, we will have a large room with several tables; at each table will be sitting a resident “specialist” in some area of interpretive methods (see below for list of participants). Anyone—grad student, faculty, other—interested in or already doing research in that area, who wants to talk with others doing the same, including people who might have specific questions about that method (e.g., I’m stuck at X, what do I do?), may approach the appropriate table, sit down, and either join the ongoing conversa-

tion (or eavesdrop on it) or wait his/her turn to discuss his/her question with the “specialist” (and anyone else who might join in the conversation).

This program has specific logistical requirements. Instead of a session “chair” there will be a session “host” (who will explain to people who come what the procedures are); instead of “presenters” we have “specialists.” (There is no separate discussant—we are all discussants!) Participation in this session will not be counted in the 2-appearance rule, as participants are “teaching” more than they are presenting their work.

Host: Cyrus Ernesto Zirakzadeh, University of Connecticut.

Specialists: Conversational interviewing: Joe Soss (University of Wisconsin-Madison); Discourse analysis: Cecelia Lynch (University of California-Irvine); Ethnography: Diane Singerman (American University), Samer Shehata (Georgetown University); Evaluative criteria: Peregrine Schwartz-Shea (University of Utah); New historical institutionalism/Science studies: Pam Brandwein (University of Texas-Dallas); Participant-observation: Dorian Warren (University of Chicago); Working with “personal” documents (autobiographies, diaries, memoirs, life histories, etc.): Lloyd I. Rudolph (University of Chicago) and Susanne Rudolph (University of Chicago); Political Theorists Doing Empirical Research: Timothy Kaufman-Osborn (Whitman College); Reflexive historical analysis: Ido Oren (University of Florida); Space/metaphor/category analyses: Dvora Yanow (Vrije Universiteit, Amsterdam).

APSA Working Group on Methodology: New Perspectives on Qualitative and Quantitative Tools August 30-September 3, 2006, Philadelphia, PA

As noted on the APSA webpage at http://www.apsanet.org/section_584.cfm, APSA is again offering opportunities for focused scholarly interaction at its annual meeting: working groups that bring together scholars who commit to attending a specified number of panels in a given area of inquiry, in combination with short courses in that same area.

For APSA 2006, our section is sponsoring a working group on methodology, covering new perspectives on qualitative and quantitative tools. Through this group the section seeks to make its panels and short courses more valuable to participants by providing continuity, coordination and shared dialogue. The working group will also convene during the APSA Meeting for a general discussion, most likely at lunch-time on Friday September 1.

APSA will provide certificates of recognition to those who participate fully in the working group. Participants in the working group will attend: (a) 2 short courses and 5 panels; or (b) 1 short course and 7 panels; or (c) 9 panels. The list of eligible panels will be made available on the APSA website and by email, and will include offerings from several different program divisions, as well as the panels and roundtables listed in this issue of the newsletter.

There is no fee to join the working group. However, since we will be meeting over lunchtime on Friday, we will be requesting that participants make a modest contribution toward the cost of their lunch.

If you are interested in taking part in the working group, please complete the form at http://www.apsanet.org/section_584.cfm. In addition, it will help us in keeping track of things if you email consortium@asu.edu to say you want to take part.

**Sixth Annual Training Institute
on Qualitative Research Methods,
Arizona State University, January 2-12, 2007**

The Consortium on Qualitative Research Methods (CQRМ) is pleased to announce the sixth Arizona State University Institute for Qualitative Research Methods for graduate students and junior faculty.

The institute seeks to enable students to create and critique methodologically sophisticated qualitative research designs, including case studies, tests of necessity or sufficiency, and narrative or interpretive work. It will explore the techniques, uses, strengths, and limitations of these methods, while emphasizing their relationships with alternative approaches. Topics include research design, concept formation, methods of structured and focused comparisons of cases, typological theory, case selection, process tracing, comparative historical analysis, congruence testing, path dependency, interpretivism, counterfactual analysis, interview and field research (including archival) techniques, necessary and sufficient conditions, fuzzy set methods, and philosophy of science issues relevant to qualitative research. Attendees will receive constructive feedback on their own qualitative research designs, and the course will also include master class discussions led by the authors of well known works which employ qualitative methods. Examples will be drawn from exemplary research in international relations, comparative politics, and American politics. The syllabus

from the fifth annual institute, available through the CQRМ webpage at <http://www.asu.edu/clas/polisci/cqrm/IQRM2006/syllabus.html>, indicates the range of the issues to be covered. Please note, however, that this syllabus will be revised for the sixth institute, and should be viewed with this in mind.

CQRМ member institutions will use their own meritocratic criteria to select students or junior faculty to attend the institute, and must notify CQRМ of their choices by October 6, 2006. Students, fellows and junior faculty not so selected, or who attend non-member organizations, may apply directly to CQRМ (see application form below). These applications must be received by October 6, 2006, and will consist of: a curriculum vitae; a list of any courses taken in qualitative or other methodology; a short (300-word) personal statement briefly summarizing their main current research project and reasons for applying to the institute; and the name and contact information for a reference who is familiar with the applicant's training and research. Applicants will be notified of the outcome by November 10, 2006.

CQRМ will cover the costs of tuition, lodging, and meals for successful applicants. Attendees will be responsible for their own transportation costs to and from Arizona State University. Participants for the institute will arrive on Monday, 1 January, and depart late Friday, 12 January, or any time on Saturday, 13 January 2007. The seminar will meet daily, beginning on Tuesday, 2 January. The final meeting is scheduled for Friday, 12 January.

**Application form for the Arizona State University Institute on Qualitative Research Methods,
organized by the Consortium on Qualitative Research Methods**

Name _____
Institutional Affiliation _____
Address _____

Telephone _____ Email _____
Discipline and Sub-field(s) _____

WHAT DO I HAVE TO SEND? (1) this form, completed, or an email with the information requested above, (2) a curriculum vitae; (3) a list of any courses taken in qualitative or other methodology; (4) a brief (300-words) personal statement summarizing your main ongoing research project and indicating why you wish to attend the institute; and, (5) the names and contact information for a reference who is familiar with your training and research. As noted below, all this information can be sent in hard copy or emailed.

WHO SHOULD COMPLETE THIS FORM? CQRМ member organizations for the 2006-2007 academic year will have spaces reserved at the institute. Members will choose their nominees using their own selection procedures. Students and/or faculty who are nominated for those slots should not use this form. Interested applicants who are: (a) from these institutions but are not so nominated; or (b) who are from non-member institutions, should use this form.

WHERE SHOULD I SEND THIS FORM? Please send application materials by October 6, 2006 to Colin Elman, Executive Director CQRМ, c/o Political Science Department, Box 873902, Arizona State University, Tempe, AZ 85287-3902. Alternatively, you can email your application to consortium@asu.edu, and you can direct any questions about the institute to this same email address.

**Studies in Comparative International Development (SCID):
Call for Contributions**

Studies in Comparative International Development (SCID) has recently moved from Berkeley to the Watson Institute for International Studies at Brown University. One of the leading journals of development studies, SCID publishes articles on issues concerning political, social, economic, and environmental change at the local, national, and international levels. In addition to original research articles on all world regions, SCID occasionally publishes reviews that summarize and assess significant, thematically linked bodies of literature and methodological essays that evaluate and/or make an original contribution to debates about the conduct of social science research. Please consider submitting your work to SCID if you have an article that fits our profile. We plan to make the review process a speedy one, so as to provide a shorter time between submission and decision. The list of scholars who make up the Editorial Collective that manages the journal at Brown, the external Editorial Board, and the guidelines for article submission can all be found at our web site: <http://watson.institute.org/ped/scid/>.

APSA-QM Section Officers

President: James Mahoney, Northwestern University
President-Elect: John Gerring, Boston University
Vice President: Nina Tannenwald, Brown University
Secretary-Treasurer: Colin Elman, Arizona State University
Newsletter Editor: John Gerring, Boston University
Division Chairs: Melani Cammett, Harvard University
Julia Lynch, University of Pennsylvania
Executive Committee: Ted Hopf, Ohio State University
Deborah Larson, University of California-Los Angeles
Gerardo Munck, University of Southern California
Charles Ragin, University of Arizona

**Qualitative
Methods**

Boston University
Department of Political Science
232 Bay State Road
Boston, MA 02215
USA

Nonprofit Org
U.S. Postage
PAID
TUSCON, AZ
Permit No. 271

Qualitative Methods is edited by John Gerring (tel: 617-353-2756, fax: 617-353-5508, e-mail: jgerring@bu.edu). The assistant editor is Joshua C. Yesnowitz (e-mail: jcyesnow@bu.edu). Published with financial assistance from the Consortium for Qualitative Research Methods (CQRM) and Boston University. Opinions do not represent the official position of CQRM or Boston University. After a one-year lag, past issues will be available to the general public on-line, free of charge, at <http://www.asu.edu/clas/polisci/cqrm/QualitativeMethodsAPSA.html>. Annual section dues are \$8.00. You may join the section on-line (<http://www.apsanet.org>) or by phone (202-483-2512). Changes of address take place automatically when members change their address with APSA. Please do not send change-of-address information to the newsletter.