Critical Comparisons in Politics and Culture

edited by

John R. Bowen and Roger Petersen

CAMBRIDGE
UNIVERSITY PRESS

Contents

Figures

1 Introduction: critical comparisons

John R. Bowen and Roger Petersen

Why compare?

The social sciences today are torn apart by a tension between two desires: to richly describe the world, showing its complexity and variability, and to robustly model the world, showing its relationships and regularities. We argue in this volume that engaging in comparisons of a few, well-understood cases reduces this tension. We offer, in effect, a case study of an encounter between two quite different disciplines, political science and anthropology. As students of society and culture, we found that we shared a stake in discovering processes and mechanisms underlying social phenomena, and that we found small-scale comparisons critical to that effort. And yet as participants in different disciplinary traditions, we continue to debate among ourselves about how best to compare, about how to interpret the patterns perceived, and about the ultimate goals of social research.

In a series of conferences and other exchanges, a collection of political scientists and anthropologists engaged in comparative study decided to put on the table what connected us and what divided us. Though a diverse lot – our objects of study run from ritual wailing to trade union disputes to agrarian transitions – we recognized in each other the dual commitment to understanding things both in their detail and in their general implications. We included no formal modelers or atheoretical monograph writers. All of us were engaged in comparative analysis of one sort or another, but some were also highly critical of much current comparative work.

We did, admittedly, approach our encounters with some fears – about disciplinary imperialisms, or about the Other's predilections for reductionism or mindless description. In truth, none of the worries have entirely been quashed, but they have been quelled, perhaps, as we have discovered that, yes, we do quite different things, and, indeed, that such is the point of the encounters. Here we wish to show and tell how these

encounters ought to enrich comparative studies for social scientists generally as they have for us as a group.

In our discussions we noted a discordance between the richness of current comparative studies in our disciplines and the narrowness of how such work is described or prescribed in handbooks and review articles. Take two recent volumes (both discussed more fully below). King, Keohane, and Verba's *Designing Social Inquiry* (1994), a masterful prescriptive text in political science, delineates a set of requirements for valid "scientific inference" that effectively reads out of social science all comparative work designed to do anything other than test (or perhaps generate) causal hypotheses. By contrast, Holy's edited *Comparative Anthropology* (1987) argues that anthropological comparisons today are designed mainly to locate culturally specific meanings, and relegates to a "positivist" past all efforts to study social regularities. The gap between these two visions could be evidence that political scientists and anthropologists have absolutely nothing to say to one another – or it could be a sign that these summations are missing something of critical importance.

We began this project betting that the latter conclusion was the correct one, and we now think we were right. We prejudiced our experiment against finding agreement by bringing together, in the same room, political scientists whose work drew on rational choice theories with anthropologists whose work was highly concerned with the culturally specific. What we found we shared was a sense that the world's complexity demands some respect even as we try to understand or isolate processes and mechanisms.

This shared commitment to describing empirical richness and accounting for it has led us to critique and try to innovate beyond current ways of doing research in our own disciplines. For example, those of us who isolate a single set of motives or interests for modeling purposes (and only some of us do that) seek to retain in the analyses the specific processes and mechanisms characterizing each case. Sometimes doing so has required creating new ways of presenting material, as in the "analytical narrative" used by Margaret Levi and her colleagues. Conversely, our descriptions are shaped by efforts to understand processes and mechanisms – how and why things got to be the way they are. This effort, too, has required new ideas, as in the thesis advanced by Fredrik Barth and Greg Urban that variation both within and across cultural boundaries should be explained by reference to similar mechanisms. In both cases we are supplementing and critiquing standard images of what strategic models or cultural accounts can be.

It is, we argue, comparison that leads us to this critical use of our

disciplinary tools – critical use, and not "application" of prefabricated tools (nor, for that matter, abandonment of social science). In this volume we show this more than tell it – we believe that exemplifying is more effective than prescribing – but we do also, here and in the other chapters, reflect explicitly on the value and limitations of particular kinds of comparisons. In design as well as in presentation, the volume is inductive, bottom-up, case-based, rather than deductive, prescriptive, law-giving. It offers the reader a set of examples to ponder, argue with, and perhaps draw from in planning comparative components for his or her own research.

Three concepts underlie these essays: comparisons, processes, and mechanisms. *Comparisons*, we argue, are at the heart of the matter for social science. We argue specifically for the value of controlled, or "small-n" comparisons of a few cases (or, as in David Laitin's chapter, a few sets of contrasting pairs). "Four plus or minus one" seems to capture what "a few" means in practice. Why comparisons, and why smallish ones, is detailed below, but the main message is that comparing several cases allows us to distinguish the important from the unimportant (or the relevant from the irrelevant, or the related from the contingent), and that limiting the number of cases allows us to deal more adequately with the complexity of social life.

We choose cases according to the questions we ask and the assumptions we make about this "complexity." When we study such "big" events as revolutions, trade union disputes, or enlistment in large standing armies (as in the projects by Margaret Levi and Miriam Golden), we may only have a few cases to start with, and the strengths and weaknesses of the analysis will depend to a great extent on the kind of information available about each (as Levi discusses).

We may decide to limit the scope of comparison to a region, or a type of society, to limit the differences between cases. This strategy of selecting closely related cases may be the result of different logics. We may, for example, be trying to control for shared features so as to isolate those elements that lead to a specific outcome, as in David Laitin's and Barbara Geddes' studies. Or we may be trying to study the variation and change in a cultural form across related societies, as in Greg Urban's and Fredrik Barth's studies. These pairs of studies start from very different questions – What are the general causal relations here? what are the specific processes of change here? – but they both depend on comparisons of closely related cases in order to find answers.[1] We may also decide to choose quite different cases so as to see if postulated relations hold up in very different contexts. David Laitin combines these two approaches; he uses the "most different case" strategy to see how

well his hypothesis holds up once he has initially tested it from a "most similar case" approach.

We use comparisons not for their own sake, but because we find that they allow us to understand better *processes and mechanisms*, the how and why, narrative and explanation, of social phenomena. Mechanisms are specific patterns of action which explain individual acts and events; when linked they form a process. As developed in political theory (see Elster 1987), they are intended to apply over a wide range of settings, and they generally refer to psychological predispositions. For example, someone might continue to keep and repair an old automobile despite the likelihood of additional costly repairs because he or she figures that a lot has already been invested in the car. This mechanism, the "tyranny of sunk costs," may also keep spouses together who would otherwise separate because they cannot accept the fact that investments in the relationship have been in vain. This mechanism is both general, in that it can be applied to a wide variety of cases (cars and spouses), and specific, in that it can be used to explain why a particular event occurs.

A mechanism approach to explanation does not, however, seek a high degree of predictive power, nor does it aim at the creation of general laws. Sometimes spouses do break up, and other mechanisms ("the grass is greener," for instance) may be at work. "If p then *sometimes* q" is the closest to a prediction that can be made within this explanatory framework. The political scientists writing in this volume by and large adopt this approach, seeking a finer-grained account of several phenomena rather than a general law. This methodological choice, sometimes associated with rational choice theory (Johnson 1996), distinguishes them from other political scientists seeking predictive power through the use of a variable approach (see King, Keohane, and Verba 1994). It also brings them closer to the anthropologists in the collection (most of whom would otherwise see little affinity between their work and that stemming from rational choice) in their emphasis on understanding particular processes rather than generating highly simplified propositions about the general relationship among two or more variables. Indeed, in his concluding chapter, Bowen argues that in all the chapters the authors make their point most convincingly when they offer microhistorical accounts of processes, and often contrasts of processes across societies, rather than static comparisons of cases.

The political scientists included here are interested fundamentally in discovering mechanisms that lead people to undertake certain courses of action under certain conditions. Margaret Levi, for example, has as her main goal understanding the mechanisms that lead people to enlist (or not enlist) in armies. But she also constructs an analytical narrative of

each country case, tracing specific macro-level pathways. Further, she tells another process story of building the model out of earlier work on taxation, looking for a very different domain against which to hone the model further, and then gradually building up knowledge of each case. (Levi thus chose her topic following a "most different case" strategy, and then compares cases of similar countries.)

For the anthropologists, both processes and mechanisms are desired objects of knowledge, but the better understanding of a particular process may be deemed more important than the uncovering of general mechanisms. Fredrik Barth's and Greg Urban's projects both involve redirecting comparative studies from the arrangement of predetermined cultural objects to the study of the processes by which forms are changed and transmitted. Ancillary to their studies, but mentioned by both as additional desiderata, is the uncovering of mechanisms that produce variation. Barth, in particular, seeks to link his fieldwork to studies of general cognitive mechanisms by which people forget and change information.

Although we find the two disciplines converging toward a renewed attention to controlled comparisons, each has its own quite distinct genealogy.

Anthropology

Anthropology exhibits continued nervousness about executing comparisons at all. When Robert Barnes (1987: 119) complains that "anthropology is permanently in crisis about the comparative method," it is the legacy of "the Comparative Method" that is at fault. This "Method" dates from the nineteenth century, and in particular from Lewis Morgan's (1871) philological studies and E. B. Tylor's (1889) cross-cultural comparisons, which he called the study of "adhesions." It is what Barth and Urban refer to as the museum approach to anthropology: isolating cultural traits and rearranging them according to such universalistic criteria as types of social structure or the relative complexity of tools.

The main traditions of American and British anthropology developed in large part as reactions to this acontextual isolation of traits. Boas and his cultural anthropology students in the United States emphasized the holistic properties of particular cultures; the founders of British social anthropology emphasized the interconnection of statuses and norms in particular societies. Yet both also engaged in comparisons of related societies or cultures. In the 1930s and 1940s Fred Eggan developed the term "controlled comparisons" to characterize studies of social variation and change in Native American societies of the southwest United States

(Eggan 1966). Regional comparisons were also used to generate and test ideas about processes, such as the rise of social stratification in the Pacific (Sahlins 1958), the development of witchcraft in Africa (Nadel 1952), or, returning to the US southwest, the development of personality through child-rearing practices (Whiting 1954). Sometimes regional comparisons were developed as contrasts, to show how different things could be along some axis within a region, as in Mead's (1935) contrasts of neighboring Melanesian societies.

Large-scale comparisons continued to be refined and expanded in the 1940s and 1950s, leading to today's "cross-cultural" method of universalistic comparisons based on a standard sample of cultures. This method typically focuses on the co-occurrence of social and cultural traits, sometimes using multiple regressions to explain the particular distribution of a feature such as "high women's status" or a certain residence rule (see Burton and White 1987).

By the 1970s and 1980s, comparative studies had been eclipsed by renewed emphases on interpretation and meaning. Large-scale cross-cultural analyses came in for particular criticism for their emphasis on traits over context, and their universalistic framework of bounded "cultures." First, critics argued, taking traits (such as "residence rules") as fixed features of cultures risks losing from the analysis many of the interesting features that good ethnography provides, including contextual determinacy (for example, residency choices depend on resources), and variation in local understandings (for example, genealogical ties to a co-resident can be reckoned in more than one way). As anthropologists turned more and more to the interpretation of local meanings, this criticism seemed increasingly telling.

Secondly, comparing across a universe of bounded, putatively independent "cultures" risks losing sight of the processes by which variation is created. The elements of a culture change over time and vary over space precisely because they have a dynamic interrelationship which can be causal and meaningful. Even in pursuit of the general hypotheses sought by practitioners of large-scale comparisons, regional variation can be a better source of data because of the control on certain variables (Mace and Pagel 1994). Although cross-cultural research has enjoyed a recent upsurge in interest, it is rarely even consulted by the majority of anthropologists; many consider it to have produced little of clear value, as recently noted by two of its major practitioners (Burton and White 1987).

Regional comparisons also have been neglected in the theoretical literature from the 1970s onward; Allen Johnson (1991) reviewed such studies and concludes that they have had little impact on the discipline

as a whole. This neglect is probably due to the combined critiques of both cross-cultural comparisons and of ethnography itself. Work labeled "post-modern" has questioned the validity of all ethnographically produced knowledge (Clifford and Marcus 1987) and has further directed theoretical discussions away from comparisons. As the editor of a volume on comparative studies put it, "the line between comparativists and non-comparativists . . . is probably more sharply drawn than ever before" (Holy 1987: 9).

And yet comparative work thrives at the heart of the discipline, particularly at the level of collaborative efforts to understand better the nature of variation and processes within regions. Controlled, regional comparisons are more widely accepted in anthropology than are universalistic ones, because they preserve a relatively high degree of contextual specificity while moving beyond the boundaries of specific societies or cultures. Much of this kind of research has been intended mainly to point out regional variations on a theme, as in a collection of studies of eastern Indonesia social organization (Fox 1980) that points to the widespread symbolic importance of the house and of the "flow of life." Similar comparativist studies of culture areas can be found for Africa (Parkin 1980), South Asia (Yalman 1967), and lowland South America (Rivière 1984). More rarely do these anthropologists identify mechanisms generating variation within the area. Barnes (1980), for example, compares marriage payments across a number of eastern Indonesian societies not only to show variations on a theme but also to argue that a causal relationship holds between the degree of trade, the levels of bridewealth demanded, and the consequent difficulty of completing payments and converting an uxorilocal marriage to a virilocal one. Mandlebaum (1988) describes the widespread ideas and practices leading to the seclusion of women across south Asia, and then explains variation in these ideas and practices by reference to women's labor participation (see also Miller 1981).

Regional analyses have been perhaps most central to studies of New Guinea societies, where they also have achieved a noteworthy theoretical sharpness. An earlier emphasis on identifying subregions by the preponderance of particular diagnostic features (for example, competitive feasting, intensive sweet-potato cultivation) has yielded to more recent studies (for example Godelier and Strathern 1990; Knauft 1993) that emphasize the ways in which different questions (for example, the spread of social organizational forms, the development of stratification) will highlight different possible configurations of subregions. Thus an initial concern with mapping of cultural forms has been succeeded by a new focus on examining the processes that generate variation (Barth

1987; Hays 1993). This change in comparative strategies is often associated with the work of Fredrik Barth, and Barth's chapter for this volume focuses on ways to study generative social processes by comparing social forms within and across cultural boundaries. Because the same set of processes may develop variation within and across cultural boundaries, this approach takes cultural forms, and not bounded cultures, as the units of analysis.

For anthropology, the emphasis in this volume on process and mechanisms recalls much of the original purpose of undertaking controlled comparisons. Eggan's studies in the US southwest were a rebuke, albeit disguised, to the scientist claims of his teacher A. R. Radcliffe-Brown that such societies had no history and that they therefore could only be understood in terms of the functional consequences of particular social forms. The comparison of neighboring societies, combined with what was known of early history, was intended precisely to reintroduce historical processes and mechanisms of change into social anthropology. What is different about today's work is in part the emphasis on comparisons of more highly interpreted realms of culture, such as the ritual wailing explored by Greg Urban and the Islamic rituals studied by John Bowen, and longitudinal analyses, as in Johnson's work. The goal of these and other analyses is understanding the processes by which cultural forms are learned, transmitted, and transformed.

Political science

Unlike anthropology, political science contains a subfield devoted to comparative studies. Arend Lijphart's much-cited 1971 article, "Comparative Politics and the Comparative Method," contains a view of the evolving role of small-scale comparisons within that subfield. Lijphart wrote (1971: 685): "If at all possible one should use the statistical (or perhaps even the experimental) method instead of the weaker comparative method." The strength of small-scale or "small-n" comparisons, Lijphart continued, lay in their ability to help create coherent hypotheses in a "first stage" of research. A statistical "second stage" would test these hypotheses "in as large a sample as possible."

Twenty-five years later, while some comparative research is conducted in the manner Lijphart recommended, much is not (see Collier 1991). In fact, the methodological coherence and division of labor envisioned by Lijphart has never developed. On the contrary, one might say that the sub-discipline of comparative politics has become either remarkably diverse or terribly fragmented, depending on one's perspective.[2]

Furthermore, as exemplified by the work and arguments of the

political scientists below, small-scale comparisons are no longer a second choice to statistical approaches, nor are they simply used to generate hypotheses as a "stage" in the research process. They are used for both theory-building and theory testing, and they form a complete research program in their own right. In order to understand the continued prominence and even resurgence of these controlled comparisons in comparative politics, it is necessary to understand both the disillusionment with other research approaches, and the innovations in small-scale comparison.

During the 1950s, political science moved away from describing the legal-formal aspects of political systems towards a more behavioralist approach. Substantively the field was dominated by the issue of "development." The Social Science Resource Council Committee on Comparative Politics became the most influential institutional actor helping to create from the late 1950s to the early 1970s a large literature on development. Many of the works produced in this era put forth universalistic typologies and chronological models: developing nations could, and would (and should), follow the Western path toward democracy with the help of institutions and processes already witnessed in the United States.

By the late 1960s, however, faith in the universalistic processes that work toward outcomes of social justice was shaken by events throughout the world. Developing countries did not follow the expected paths, and Vietnam was a disaster. The last great grand synthesis of the field, Huntington's *Political Order in Changing Societies* (1968), reflected the original developmentalists' loss of optimism. The Social Science Resource Council Committee on Comparative Politics was disbanded. The backlash against the developmentalists produced a whole new set of general models. Dependency theory, corporatism, and bureaucratic-authoritarianism are the most well-known and direct responses to the perceived failures of the developmentalist approach.

However, these general models proved inadequate in explaining the complexity of modern politics: Asian newly industrialized countries (NICs) produce booming economies while other developing economies flounder; military regimes fade from Latin America while fundamentalist revolutionary regimes appear elsewhere; communist regimes fall but former communists win elections; mass ethnic killing in Rwanda and the former Yugoslavia occur simultaneously with peaceful change in South Africa and the Middle East; and so on. As complexity increased, two dominant approaches, model-building at the level of grand theory and large-scale statistical studies, went into relative decline.

The focus of the political comparativists in this volume is less on

sweeping general models and more on explaining better-defined phenomena. Miriam Golden explains a set of labor actions in industrialized states; Barbara Geddes explains bureaucratic reform in Latin America; David Laitin isolates a set of conditions explaining nationalist violence; Margaret Levi's work explains variation in conscription policies and responses in several Western states. Following William Riker (1990), Golden describes her choice of topic and scope by asserting that "a narrow focus to attain a proper solution is a better research strategy than a broad focus that fails to generate conclusive results. By narrowing the focus of the phenomena under study, we reduce the trade-off between analytic rigor and empirical accuracy." An increasing number of comparativists have come to agree with this argument.

While large-scale studies are still prevalent in comparative politics, faith in cross-cultural and cross-national statistical study has diminished with increased awareness of problems associated with conceptual "stretching," unreliable measures, and improper specification of domain and units.[3] As Sartori (1970) has pointed out, the very concepts used to define independent and dependent variables often translate across societies only with the greatest difficulty. As more cases are included in a given study, the basic concepts are often "stretched" to incorporate them, sometimes to the point of meaninglessness. Furthermore, heightened appreciation of cultural difference has generated skepticism of statistical measures. For instance, does the gap between expected income and actual income really measure relative deprivation in both France and Indonesia? Does "income" have the same meaning and relevance in both societies? When does the social scientist know which cases belong in the sample if knowledge of cases is superficial (as in most large-scale studies)?

In addition to some of the more intractable methodological problems involved with large-scale statistical studies, some scholars are not satisfied with the very nature of the explanation that such work provides. Rather than simply identifying probabilistic relationships between sets of variables, many comparativists would rather work to identify the nature of causal linkages among parts of a process. The work of David Laitin (see chapter 2) comprises such an effort.

Many of today's political comparativists are skeptical of the abilities of general models and large-scale statistical work to capture the complexity of their subject matter; however, they remain committed to social science methods that allow for generality. Skepticism has not produced the desire to do purely descriptive and highly specific work. Margaret Levi speaks for many comparativists when she writes in chapter 8 that "an overemphasis on specificity . . . obscures the commonality among

cases and places." It is at this point that small-scale controlled comparison comes into play. Through a focus on process and mechanism within the detailed study of the cases, much of the complexity of political life can be addressed while maintaining an ability to generalize. Through a focus on control, the benefits of social science logic (for example, covariation and falsifiability) are preserved.

Despite fragmentation in actual practice, there is a political science tradition of attempting to delineate one fundamental logic that underlies all comparative study, both quantitative and qualitative, and perhaps all of social science. In their influential 1970 work, Adam Przeworski and Henry Teune (1970: 86) conclude: "Although the phenomena under consideration vary from discipline to discipline, the logic of scientific inquiry is the same for all social sciences. As the theories explaining social events become general, the explanations of particular events will cut across presently accepted borders of particular disciplines." In another influential book published nearly two and a half decades later, Gary King, Robert Keohane, and Sidney Verba (1994: 4) write: "A major purpose of this book is to show that the differences between the quantitative and qualitative traditions are only stylistic and are methodologically and substantively unimportant. All good research can be understood – indeed, is best understood – to derive from the underlying logic of inference."[4] Charles Ragin's work on comparative method (1987) sees differences between variable-oriented and case-oriented methods but believes these differences are reconcilable. He proposes a synthetic approach employing Boolean algebra.

Our approach differs. While we applaud the search for common ground, we believe that the differences among the disciplines are more than a matter of style. Certainly, the prevailing goals vary among fields, if not the respective logics. Rather than trying to convince social science practitioners that there is one underlying logic, or developing a new synthesis, we believe that interdisciplinary progress might best be made by presenting choices and trade-offs made in the course of quite distinct research projects. We believe that knowing a wider range of possible ways of comparing will both help individual researchers in their own work and help build bridges across disciplines.

Although not its central concern, this volume thus indirectly addresses the notion of "one social science logic" that seems to preoccupy some political scientists. If both political scientists and anthropologists choose similar strategies when confronted with similar dilemmas of comparison, despite their different rhetorics and style, support for the idea of one pervasive logic is provided. On the other hand, if substantial and consistently different research choices are made, then the outlines

of different logics of inquiry emerge. We will let the reader be the judge on this issue.

Validity and generality

Despite the convergence of these two fields on small-scale comparisons, real contrasts emerged from our discussions about what political scientists and anthropologists wish to emphasize. Both disciplines encourage their practitioners to develop ideas about how the world works that are faithful to those workings and that also have some degree of generality. But in our discussions the anthropologists tended to emphasize above all the validity of their knowledge, and the political scientists the value of constructing a model that can explain more than one case.

By "validity" we mean the degree to which the account of something picks up processes, ideas, or relationships that are indeed there in the world. Insisting on "validity" does not imply a correspondence theory of truth (that a true description maps one-to-one onto the world), but only that some descriptions are better than others, and that the kinds of things anthropologists do when in the field – checking with many people, listening in on discussions, and living through events – are particularly good ways to arrive at a good description. (We do not intend the statistical meaning of "validity.")

Greg Urban emphasizes that what appears to be a "simple" description of a cultural form already requires several levels of comparative activity, for members as well as observers of a society. Urban began his study of ritual wailing from his puzzlement that what seemed to be crying was performed in contexts of welcoming someone home. He then tried to understand the behavior he had seen by comparing different instances of it in the Amazonian society where he was living, looking for when it is performed and what meanings people seem to be imputing to it. He points out that this process of comparing within a culture is precisely what children in the society do when learning their culture. Now, Urban could have stopped there and simply reported how things worked in the one society. But his interest is mainly in the processes by which culture is transmitted within and across boundaries, so he began to look for similar phenomena elsewhere. The problem arose that in other, related societies what he might have taken to be a central component of the wailing, for example the "cry break" or creaky voice, were missing or difficult to hear, even though the forms were used in the same way. Urban concludes that we cannot begin with a singly defined phenomenon and then see if a society has it or not, because the form changes as it is transmitted. Instead, he urges that we take the processes

of transmitting forms (and generating variation in them) as our object of study. The most valid knowledge of the cultural form even in one society, then, already takes into account the possible directions of its change.

Fredrik Barth joins Urban in attributing some of what they see as comparativist mistakes in anthropology to the close historical relationship of the discipline to the ethnographic museum, where things often were, and are, laid out according to function. This layout implies the prior assumption that the objects, despite their differing cultural contexts, are essentially the same. Such thinking, write Urban and Barth, has led some anthropologists to treat ideas and behavior similarly – as things that can be classified according to one-word rules, forms, or functions.

Barth questions the usefulness of comparing either cultures or "traits." He offers two arguments. First, that isolating cultural traits – a rule about whom one marries, or the general status of women – and subjecting them to cross-cultural statistical analysis omits their context-specific character and may systematically, not just randomly, distort the analysis. To take an example from another prominent anthropologist, Pierre Bourdieu (1972) argues that members of Kabyle society in Algeria say that a man ought to marry a cousin related to him through other males, and this may be coded as the preferred marriage for comparative analyses, but this trumpeted "rule" disguises the fact that many couples are brought together through female ties; men then reinterpret the marriage to highlight (often more distant) male ties.

Secondly, Barth argues that the most important mechanisms generating variation may operate across societal boundaries, and thus are not definitive of specific cultures or societies. It is artificial, he writes, to distinguish between the variation one finds across societies from that which one finds within a single place if both are generated by similar mechanisms.

One could suggest that part of his argument follows the logic of comparisons set out by John Stuart Mill. Thus, Barth writes that studying the diversity within the larger New Guinea area in which he worked helped him to focus on those features of Baktaman ritual that were "foundational," as he puts it, and those features that were due to "the flux of free variation." This moment of his analysis is a variant of Mill's analysis of differences: eliminating those elements which are not found in a number of closely related cases (Zelditch 1971). Barth also looked for covariation between elements of ritual and elements of context, within Baktaman society or across societies. But his goal is not to use such a method to explain variation, but rather to differentiate

elements that vary from elements that do not, and then explain variation by studying the processes through which people learn, transmit, and alter knowledge.[5]

Barth and Urban highlight the importance of studying variation and processes of transmission within and across societies as a comparative study. Their chapters raise the following question about theory and objects of study: to what extent is their position (for variation, against comparing bounded units) a function of the kind of cultural objects they study? Ritual form may be inherently less amenable to correlational analyses than, say, local-level conflict. On the other hand, studying the transmission of discourses of ethnic identity and nationalism is hardly far-fetched, as Benedict Anderson's *Imagined Communities* (1983) exemplifies.

Barth's final point suggests one possible way of relating studies of variation to other forms of analysis. He cautions that *descriptions* of variation do not lead directly to *explanations* of variation. Explaining variation requires us to draw on hypotheses about why people act or think in certain ways in general: for example, what cognitive processes might have generated variation in ritual, or how differences in political structure might have produced variation in Balinese village form.

Miriam Golden's analysis of trade union disputes began, like Urban's study, with an anomaly. After years of studying trade unions in Italy, she puzzled over the fact that unions called strikes that were virtually unwinnable, in that the stated goal of the strike, preventing job loss, clearly could not be reached. Why this apparently irrational behavior? she wondered. She then compared decisions to call strikes in several industrialized countries, and found a second anomaly. Although union leaders say they strike to prevent downsizing, they do not seem to respond more forcefully when more jobs are at risk. She concluded that the real motivation behind strikes was protecting the union itself. The Italian strikes were realistically conceived; they failed because leaders overestimated employees' willingness to follow their strike call. Golden's argument is convincing precisely because she attends to the details of process in each of her cases: what leaders and followers knew, how they assessed their chances, what happened after the strike call.

Although both use comparative strategies to better understand basic processes and mechanisms, Urban and Golden follow very different logics of research. Urban's question is fundamentally about cultural processes, although he studies mechanisms of transmission and learning, whereas Golden's is fundamentally about mechanisms that shape decisions, although she studies processes of union formation. One can imagine a series of Urban-like questions asked of Golden: how did

union leaders learn what ought to be fought for and what did not matter? in what ways did ideas and norms about strikes, leadership, wages, and so forth spread across these societies (through pamphlets and books, congresses and visits)? One can also imagine a series of Golden-like questions asked of Urban: what leads someone to initiate wailing? are there risks or costs if one fails to wail at the correct time or in the correct way? (These questions may be more interesting to Urban than the questions anthropologists usually expect to hear from other social scientists, such as: under what conditions would one expect to see wailing associated with welcome across cultures? In fact, that question is more likely to emerge from the large-n tradition of cross-cultural research in anthropology than from the political scientists included here.)

Golden explicitly rejects the idea that the best way to discover union leaders' beliefs, knowledge, and intentions is to ask the leaders directly. Because she believes that they will systematically distort their answers to such questions, she decided to build a model of their actions based on comparative data and inferences from a number of strike decisions, rather than from in-depth interviews. In our discussions some of us disagreed with this decision on grounds that, if systematic, differences between leaders' statements about their own actions and Golden's inferences could be quite interesting. We share the goal of discovering what union leaders know and intend to do; we differed, and continue to differ about how important actors' self-reports, rhetorics, and debates are to the analysis.

Generality is the second desiderata that guides our work. By "generality" we mean the capacity of an idea or hypothesis to account for a number of cases. (The political scientists tended to favor the term "robustness" to refer to the narrower concept of a model's capacity to explain a number of cases.) As Urban points out, even understanding the meaning of a single cultural form in one small-scale society requires a degree of generality, in that the hypothesis about what the form means must stand up over a number of its occurrences. Such hypotheses, whether about meaning, decisions, or social processes, imbue an account with some degree of generality, in that the account concerns more than a single event.

The nature of the material also shapes the meaning that "generality" has in these contributions. John Bowen analyzes variation in local forms of religious practice, when all practitioners develop local rituals with many of the same Islamic prescriptive texts in mind. Variation is within a controlled domain, and one would expect certain elements (such as the efficacy of sacrifice, or the importance of self-abnegation) to appear in

many societies. Certain general associations might appear, although the analysis is mainly intended to explore a field of variation – as in the analyses by Urban and Barth (on New Guinea) of local ritual knowledge. But in the latter cases ritual knowledge is not encompassed by textual or other authority and thus can change from one case to the next, in the manner of "family resemblances," until endpoint cases share no features. The relatively controlled nature of the Islamic prescriptions makes possible a series of successive contrasts, within Gayo society, within the larger province, and finally between Indonesia and Morocco.

All of us find ourselves tacking back and forth between model building and the interpretation of data, but we differentially highlight particular moments in that movement. Some of us stress the way in which new observations lead to new theoretical understandings, as when Urban's cried welcome led to a new set of ideas about cultural processes, or Johnson's return visit to his Brazilian research site led him to revise his understanding of how political-economic changes reshape worker consciousness.

Like many projects, Johnson's research was begun from a combination of personal experience, theoretical training, and the lack of correspondence between the two. Having been trained to think of peasant consciousness as class consciousness, Johnson was surprised to find peasants thinking primarily in terms of patron-client networks during his fieldwork on Boa Ventura in Brazil. He developed a model of peasant behavior based on "client consciousness," as opposed to a view of "proletarian consciousness" dominant in the literature of the time. But he also wished to explain the conditions under which client consciousness would become proletarian consciousness. In order to accomplish this task, he returned to Boa Ventura twenty-two years after his original research. Class consciousness had not developed, he found, largely because of state welfare interventions. The comparison thus usefully points to limitations of models that take insufficient account of the various policies that states might pursue. Johnson then compares the Boa Ventura case with transitions from patron-client systems to more capitalistic ones elsewhere in Latin America, in countries where state intervention was less beneficial to peasants. He finds, as his revised model would predict, peasant revolt and the development of class consciousness. His model thus employs two stages of comparison: longitudinal, to develop a hypothesis about change (which can also be seen as experimental, a sub-genre of comparison), and controlled, regional comparisons, to check this hypothesis against other cases, noting especially cases where the major intervening variable (state welfare policy) was present or absent.

Others of us underscore the way in which an observed set of differences in the world could be explained by a new model of social action, as when Golden is able to explain the differential likelihood of strike calls by a model emphasizing union leaders' perceptions of the current threats to union strength, or when Laitin accounts for different degrees of violence.

In order to explain the use of violence in some cases of national revival but not others, David Laitin begins with a paired contrast, between Catalonia and Basque Country. He builds an explanatory model, whose critical mechanism is the tipping point where sufficiently many people participate in the movement to make the costs of participation drop. He then looks to explain why it is more difficult for nationalist leaders to recruit followers in some cases than others, finding answers in such factors as social networks and language histories. When these tipping points are more difficult to reach, leaders sometimes choose a violent course of action in order to raise the ante of sticking with the status quo and push more people toward commitment. Laitin then moves from the Catalonia/Basque Country contrast to one between Ukraine and Georgia, to enhance the plausibility of the account.

Laitin calls this kind of research emphases "dependent variable driven," in that observed differences in the world are the catalyst for a research project. Laitin agrees with Barth that comparisons need to be complemented by general hypotheses before one can claim to have explained differences.

Some of us place still more emphasis on the process of constructing and testing a model – but some of these model-driven projects also began with an intriguing puzzle. Barbara Geddes started with knowledge of Brazilian politics and society. She was struck by Brazil's relatively high growth rate and successful civil service reform. Her curiosity about Brazil's unusual path compared with those followed by other Latin American countries led her to ask a general question: under what conditions is civil service reform enacted in Latin America? Like Johnson, Geddes found a dominant paradigm, in her case that of collective action, insufficient to explain the case at hand. She then fashioned a model that reflected real Brazilian social-political life. Her model compares politicians' attitudes toward civil service reform by whether they are members of the dominant party or not, and works well for Brazil.

At this point Geddes had accounted for the Brazilian case, but because the model had been built to fit that case it did not yet test a general hypothesis. The pay-off matrix of the game did suggest that if a country had political parties of equal size, individual politicians, seeking

greater electoral support, might support civil service reform. Geddes thus selected the four other Latin American countries with sufficiently long periods of democracy on which to try out this proposition. Her selection controlled to some degree for the Iberian political culture of the region and (because the four countries all were in a middle income range) for gross differences in economic development. A further set of contrasts was provided by changes in some of these countries from equal-size party situations to dominance by one party, a change that led to a scaling back of reform, as the model predicted.

Running through all these quite distinct research histories is the necessity of comparison to check an initial understanding against a broader field of data. This move toward comparison is not a specific method, but a necessary logical part of any research process that we can imagine. Comparisons, it should be remembered, are part of how we understand any social phenomenon, because we necessarily compare different utterances of what may or may not be the "same" expression, or different occurrences of what may or may not be the "same" ritual. They are part of how we account for any set of events, because we necessarily wish to specify which values of one or more variables lead to a certain occurrence and which do not.

In the end we see in our own work trade-offs between generality and validity, but agree that we must understand social life on the micro-level, in terms of how people come to know and act as they do. We disagree as to how best to do this. Some of us adopt a relatively thin view of intentions in modeling, others start with a thicker view. Some believe that actors' statements are necessary data; others disagree. But we see these decisions as strategic ones for purposes of obtaining particular kinds of results, and not as statements of ontological positions. Indeed, we would argue that creating analytical models not only is compatible with fieldwork or historical research but ultimately *requires* it for gathering sufficient information about strategies, decisions, and beliefs.

But we also emphasize that the fullest understanding of these events will incorporate the larger structure within which events and institutions are embedded. In our discussions, William Skinner (whose paper was not, in the end, included) pointed out that if you control for macro-region, apparently contradictory theories can often be resolved. Regarding the ongoing debates about peasant economic motivations, for example, debates often associated with the work of James Scott and Samuel Popkin, we might regularly find "maximizing peasants" in certain spaces of a macro-structure, and peasants focusing on minimizing risk and ensuring overall survival in other such spaces. David Laitin commented that macro-structural analysis, when fully elaborated,

should be considered as fundamental in any attempt at a controlled comparison. If such data were available, suggested Laitin, and it showed that his cases differed on the macro-dimension – for example, Catalonia is a regional system with a city at its center, while Basque Country is not, with the expectation that the former would be more integrated – this factor would become the basis for new hypotheses and better controls for future comparisons.

"Critical comparisons," we believe, can best be made when, as we are engaged in our discipline-specific work of analyzing, modeling, comparing, we keep in the backs of our minds other possible strategies, other versions of social science. Our goal for this volume is to provide a set of cases that can be held in mind, some in the forefront, some in reserve, as reminders of other courses we might wish to take.

Notes

1 Although we concentrate on small-n comparisons, we recognize the value of other research designs, including both large-scale comparisons and, at the other end of the continuum, the use of a single case in "crucial case" designs. In the latter approach, the researcher analyzes one case judged to be the most likely case to fit the theory. Failure to fit then leads the researcher to abandon or modify the theory. We emphasize comparisons of multiple cases because we find them better able to generate new theory and shed light on already developed theories (see Eckstein 1975; Rogowski 1995).

2 For generally positive views, see Verba (1991) and Wiarda (1991). Both mention the "islands of theory" concept of comparative politics first suggested by Stanley Hoffman: "He [Hoffman] argued cogently and convincingly that, because the comparative politics field had lost its earlier unity, those active in the field should now accept this fact realistically rather than simply lamenting it or wishing it away. Rather, what we have now are various 'islands of theory' appropriate for the several, quasi-self-contained parts of the field – political culture studies, political socialization studies, political party studies, interest group studies, political economy studies, voting behavior studies, public policy studies, government performance and effectiveness studies, and the like" (Wiarda 1991: 245). See Almond (1990) for a collection of essays on the divisions within comparative politics and political science on the whole.

3 We do not wish to be overly negative about cross-national statistical models in political science. These models continue to produce fine statistical work, particularly on party and electoral systems, and may be most appropriate for understanding such issues as the relationship between GNP and the stability of democracy, where multiple mechanisms are in play.

4 Five critical reviews of this work are found in "The Qualitative-Quantitative Disputation: Gary King, Robert O. Keohane, and Sidney Verba's Designing Social Inquiry: Scientific Inference in Qualitative Research," *American Political Science Review*, 89: 454–481.

5 Stanley Lieberson (1991) argues that, taken alone, the Mill method cannot provide adequate accounts of why anything happened, in part because variables that will be dismissed as causal (because, say, present in two contrasted cases) may none the less play a causal role, and in part because the interaction of variables is not captured.