

APSA-CP

Newsletter of the APSA Organized Section in Comparative Politics

Volume 9, Number 1

Pps. 1-2, 4-5.

Winter 1998

Letter from the President Comparative Method in the 1990s

David Collier

University of California, Berkeley – dcollier@socs.berkeley.edu

The late 1960s and early 1970s saw the emergence of a new literature that established the “comparative method” as a fundamental component of the comparative politics enterprise. Comparative method was viewed as the systematic analysis of a relatively small number of cases (i.e., a “small *n*”), and was understood in contrast to the statistical, experimental, and case-study methods.

A quarter of a century later, we are now in the midst of a major new round of debates on this branch of methodology, and I wish to use my first letter from the president to make some observations about these debates. I focus here on what may be thought of as the division of labor in comparative politics between the comparative method and the statistical method, and also on the issue of conceptual validity, a long-standing concern of the comparative method. I will refer in my discussion to six articles in this issue of the *Newsletter* that reflect important facets of these debates.

Comparative Method vis-à-vis Statistical Method

How should we understand the role of the comparative method in relation to the statistical method? One view was offered in Arend Lijphart’s seminal article on “Comparative Politics and the Comparative Method” (*APSR*, 1971). Lijphart in effect saw the comparative method as a way station, at which analysts may stop to carry out initial tests of important hypotheses. Later, after scholars have done the hard work to create more sophisticated data sets, they should move on to research designs based on stronger empirical tests, utilizing the statistical method. According to this initial formulation of Lijphart’s view, the comparative method should play an important, but perhaps transitional, role within any given substantive area of research.

Given that many scholars believe that the statistical method is “obviously” a stronger approach, it is important to emphasize that Lijphart subsequently called attention to strengths and weaknesses of both the comparative and the statistical method. He underscored, among other things, the advantages of the comparative method in dealing with problems of conceptual validity, suggesting that perhaps we need to think of the comparative method as more than just a way station.

In that spirit, I view the comparative method as an important approach in its own right, one that is not limited to transitional or exploratory work. Within the field

of comparative politics, it remains a central methodology which scholars employ to accomplish important analytic tasks, and to which they periodically return, even at more “advanced” stages of research.

The cycle of returning to the comparative method takes various forms. First, it can be seen in the evolution of research on specific substantive topics. In a given area of study, a phase of research based on statistical analysis may be *followed*, rather than *preceded*, by a phase in which small-*n* comparison adds crucial insights. Scholars routinely go back to a small number of cases to assess the validity of conceptualization and measurement, as well as to refine causal inferences. Thus, small-*n* analysis has an important role to play, even when data for large-*n* studies are available.

A recent example of this sequence is found in the democratic peace literature, which analyzes the apparent tendency of democratic countries to go to war less frequently, at least with one another. The Bennett and George article below argues that an initial phase in this literature based on statistical analysis has been complemented by subsequent work in the comparative case-study tradition. Another example is found in the literature on the political economy of advanced industrial societies, in which a central goal has been to evaluate political explanations of national economic performance. In these studies, following an expansion of the *n* and a shift to more complex statistical modeling based on pooled time-series cross-section data, concern has subsequently been expressed about the reliability of causal inferences drawn from this type of data. One possible route to follow in light of

this concern is a new iteration of small-*n* research.

The recurring importance of the comparative method is also evident in the trajectory of methodological discussions. In debates of the 1990s on the relationship between quantitative and qualitative research, scholars have repeatedly gone back to insights drawn from the comparative method. The contributions below by Charles Ragin, John Stephens, and Timothy McKeown reflect these debates. Ragin compares the approach to causal assessment adopted by the comparative method with that of the statistical method. He highlights the problem of establishing “sufficient” causes and argues that this type of causation is more effectively analyzed by a new approach to the comparative method – based on “fuzzy logic” – than by statistical analysis. Stephens shows how the comparative method and the statistical method deal with the small-*n* problem, Galton’s problem, and the “black box” problem, offering the interesting observation that these two methods can suffer from similar dilemmas of indeterminacy in causal inference. McKeown adopts a different point of departure within the spectrum of methodologies, focusing on how causal inferences can be constructed on the basis of evidence and hypotheses derived from a single case. He contrasts this case-based approach with the statistical approach to causal inference, and his contribution serves as a useful reminder of the degree to which comparative work ultimately rests on the meticulous interpretation of individual cases.

A return to the comparative method is likewise seen in the trajectory some

(continued on page 4)

(Collier, continued from page 2)

times followed by specific research projects undertaken by individual scholars. Within a given study, a scholar focused on a small number of cases – for example, a limited number of national political regimes – may supplement the small-*n* comparative analysis of national units with further analysis focused *within* each country, based on a large *n*. Such within-case assessment might involve, for instance, analysis of public opinion data, national budgets, or other kinds of within-nation data that entail a large number of observations. However, to the extent that the goal is to bring explanatory insights from the within-case analysis back up to the level of the national political regimes that were the initial focus of concern, this ultimately remains a small-*N* analysis. Hence, the scholar will return to the comparative method in the final stage of the study.

Finally, the recurring importance of the comparative method is evident not only among scholars pursuing alternative methodologies, but also among analysts using diverse theoretical tools. For example, the forthcoming book summarized below by Peterson and Bowen includes five chapters in which game theorists test their models using carefully executed small-*n* comparisons.

To summarize, one sees not only periodic movement away from the comparative method, but also periodic movement back to it. Let me explore this theme further with reference to the issue of conceptual validity.

Conceptual Validity

Conceptual validity is an abiding issue in comparative research. The concern with validity is animated in part by a recognition of the trade-off between 1) the drive to extend our theories and hypotheses to a larger number of cases, and 2) the problem that if we extend them too far, conceptual stretching may occur, in that our concepts no longer validly fit our observations. This concern likewise derives from a fundamental preoccupation of many small-*n* analysts:

they worry that indicators employed in large-*n* cross-national research frequently fail to measure the concepts they purport to measure. Whatever vision one may have of the “scientific” status of comparative politics, this vision must include a central concern with validity. A focus on conceptual validity, correspondingly, has a prominent place in writing on comparative method. Major statements in the 1970s include Sartori’s analysis of conceptual stretching in “Concept Misformation in Comparative Politics” (*APSR*, 1970), and Przeworski and Teune’s recommendations in *The Logic of Comparative Social Inquiry* (Wiley, 1970) for adapting measurement to specific contexts, including potentially the use of what they call system-specific indicators.

Recent work has refined these perspectives in several ways. Charles Ragin has developed an analysis that parallels Sartori’s discussion of the intension (meaning) and extension (domain of relevant cases) of concepts. Ragin introduces the label “double fitting” to characterize the process of mutual adjustment between these two dimensions that often occurs in the course of concept formation. Shifts in meaning (i.e., in the definition of the concept) can push the analyst to adjust the corresponding domain of cases, and shifts in the domain of cases can necessitate an adjustment in the meaning, so as to maintain conceptual validity. Ragin suggests that in much research, as this double fitting proceeds, the domain of cases under investigation may remain fluid during initial phases of a study. Thus, in a comparative study of revolution, shifts in the definition of the main concept can dramatically change the relevant domain of positive and negative cases. Such shifts likewise occur in the broader evolution of scholarly research programs.

Given that establishing the domain of relevant cases is an essential underpinning for addressing various methodological issues, it is productive to recognize that this initial fluidity in defining this domain does indeed occur in many studies. It is impossible, for example, to make

judgements about selection bias until the domain of cases is established. A warning about another kind of bias is also essential. This process of double fitting should be used *appropriately* to refine concepts, and not *inappropriately* to come up with a set of cases that conveniently confirms the researcher’s preferred hypothesis.

A further contribution by Ragin to the discussion of validity is summarized in his article below. In a notable departure from his earlier focus on the dichotomous variables employed in Boolean algebra, he explores the possibility that the logic of fuzzy sets may sometimes offer a more valid operationalization of our concepts than does either dichotomous or quantitative measurement.

Another aspect of validity, linked to the idea of system-specific indicators, is explored below by Locke and Thelen. Whereas system-specific indicators were originally proposed as an approach to *quantitative* comparison, these authors suggest that scholars conducting *qualitative* research at times must engage in a parallel process of “contextualized comparison.” Thus, to generate conceptually equivalent observations in relation to a given concept, it is sometimes necessary to focus on what at a concrete level might be seen as distinct types of phenomena. For example, scholars who study national responses to external pressure for economic decentralization and flexibilization are sometimes concerned with identifying analytically equivalent “sticking points” where sharp conflicts emerge over this economic transformation. In the domain of labor politics such conflicts may, in different countries, arise over wage equity, hours of employment, work-force reduction, or shop-floor reorganization. The scholar must look at these different domains to make analytically equivalent comparisons that correspond to the concept of “sticking point.” Similarly, in *Shaping the Political Arena* (Princeton, 1991), Ruth Berins Collier and I applied the concept of the “initial incorporation” of the labor

movement in a parallel manner, recognizing that analytically equivalent observations linked to this concept entailed, in concrete terms, somewhat distinct phenomena in different countries.

Given the prominence of Przeworski and Teune's proposal for system-specific indicators, it is curious that in the intervening years this approach has not been used more frequently. Locke and Thelen's examples of comparing "privatization" and "globalization" across the countries of Eastern Europe help to clarify this puzzle. These examples suggest that comparativists who are closely familiar with the contexts they are comparing may in fact routinely employ this approach of contextualized comparison. Yet they often do so instinctively, rather than self-consciously. Following the phrase of Molière, it could be said that comparativists are sometimes "speaking prose" without recognizing it – i.e., carrying out contextualized comparison without being explicit about it. Clearly, it is preferable to make this practice explicit, and the Locke and Thelen article should help push scholars to do so.

Effective use of double fitting and contextualized comparison requires careful attention to the structure of concepts, to how concepts embody meaning, and to how scholars can most effectively use concepts in pursuit of their analytic goals. The recent small-*n* and case study literature on democratization offers examples of both successes and failures in the use of concepts. These successes and failures arise in part out of scholars' responses to two conceptual challenges posed by the recent world-wide wave of democratization. Analysts seek both to increase analytic differentiation in order to capture the diverse forms of democracy that have emerged, and also to avoid the conceptual stretching which arises when the concept of democracy is applied to cases for which, by relevant scholarly standards, it is not fully appropriate. A dilemma arises from the fact that efforts to increase differentiation through introducing finer distinctions may produce analytic categories that are more vul-

nerable to conceptual stretching.

Analysts have fine tuned their concepts in many different ways as they pursue these contending objectives, including the creation of what may be called "diminished" subtypes of democracy. For example, the concept of "illiberal democracy" can serve to differentiate cases where the protection of civil liberties is seen as inadequate; and because it is a diminished subtype, it avoids conceptual stretching by specifically *not* making the claim that these are full instances of democracy, which by standard definitions they clearly are not. In the hands of careful, well-disciplined scholars, such conceptual innovations can yield better research.

However, this proliferation of conceptual forms also has a down side. For example, the literature on democratization has spun out literally hundreds of democratic subtypes, and too often these subtypes either are not clearly defined, or are not employed in a consistent manner, or both. Consequently, any gains that might be achieved in finer analytic differentiation and/or improved conceptual validity may be cancelled out by the resulting conceptual confusion. When such confusion arises, it is essential for scholars to engage in a self-conscious, critical evaluation that systematically appraises existing usage of concepts and seeks to channel it in more productive directions.

Researchers who work closely with a small *n* are supposed to have the advantage of "knowing their cases," thereby helping them to avoid the problems of validity that may arise for scholars who are not as familiar with the contexts they are studying. Yet in addition to knowing their cases, scholars need a disciplined understanding of how to employ concepts, along with a firm grasp of how to organize concepts into worthwhile theoretical arguments. The challenges of learning and teaching these skills, as well as applying them effectively in different substantive domains of research, must be an abiding concern in the field of comparative method.

References

- Nathaniel Beck and Jonathan N. Katz, "What To Do (and Not To Do) with Time-Series Cross-Section Data," *American Political Science Review* 89, No. 3 (September 1995): 634-647.
- David Collier and Steven Levitsky, "Democracy with Adjectives: Conceptual Innovation in Comparative Research," *World Politics* 49, No. 3 (April 1997): 430-451.
- Bernhard Kittel, "Sense and Sensitiveness in Pooled Analysis of Political Data," *European Journal of Political Research*, forthcoming, 1998.
- Arend Lijphart, "The Comparable-Cases Strategy in Comparative Research," *Comparative Political Studies* 8, No. 2 (July 1975): 169-171.
- Charles C. Ragin, *Constructing Social Research: The Unity and Diversity of Method* (Thousand Oaks: Pine Forge Press, 1994), chap. 4.