Newsletter APSA - CP

Volume 15, Issue 2
Summer 2004

The Organized Section in Comparative Politics of the American Political Science Association

Contact
Decio Hall, Box "D"
University of Notre Dame
Notre Dame, Indiana 46556
Tel. 574-631-5681
Electronic correspondence preferred
http://www.nd.edu/~apsacp

Letter from the President

Beyond the Comparative Method

Peter A. Hall
Harvard University
phall@latte.harvard.edu.

There is a paradox at the heart of comparative politics. The field is closely associated with a specific method—the "comparative method" inspired by Mill, elegantly defined by Lijphart (1971; 1975a) and elaborated in many texts (cf. Smelser 1976; Collier 1991). When I was a graduate student, this method was presented as virtually synonymous with the field. But even its most prominent exponents admit that it provides a weak basis for assessing the causal claims crucial to political explanation. "If at all possible," noted Lijphart (1971: 685), "one should generally use the statistical (or perhaps even the experimental) method instead of the weaker comparative method."

The limitations of the comparative method are well known. As conventionally defined, it is oriented to a particular type of explanation focused on finding a few key variables said to be the principal causes of a specific kind of outcome. Like the statistical method, it is essentially correlational, assessing the validity of an explanation by inspecting the correspondence between the values of the explanatory variables and those of the dependent variable or outcome. What distinguishes it from the statistical method is the small number of cases examined.

With few cases, however, it is difficult to secure wide ranges of variation on all potentially relevant variables. For many years, the problem of "too many variables, too few cases" obsessed the field. Ingenious attempts were made to control for the effects of other variables, by selecting cases said to be similar in all relevant respects except on the outcome and explanatory variables of interest or cases different in all respects except for a correspondence between the values of the explanatory variables and outcomes (cf. Przeworski and Teune 1970). Scholars scoured the world for the "critical cases" in which, on current knowledge, their findings could be presented as least likely to hold and hence likely to be valid elsewhere (cf. Eckstein 1975). But, since the "laws" of social science were often thought to be probabilistic, one could never be sure whether one was finding the relevant regularities or encountering the exceptions to them. Despite the prominence of the comparative method, we all knew we were swimming upstream.

How then is the success of comparative politics to be explained? Few would deny that the field has vastly increased our knowledge of the world. No respectable department of political science would be without scholars of comparative politics. There is much to be learned from the leading journals of the field and causal explanation is the currency in which they deal. How

(Continued on page 2)

Section Officers

(Continued from page 1)

can a field so closely associated with such a weak methodology have accomplished so much?

On reflection, I think the answer is that relatively few scholars of comparative politics have ever relied heavily on the comparative method, at least as conventionally defined. Instead, the success of the field rests largely on the ways it has moved beyond that method. Four steps, taken by many scholars, seem important.

The first step was to broaden the object of inquiry. Instead of establishing the causal impact of a few key variables, many comparativists have seen their task as one of mobilizing the power of comparative inquiry to develop new theoretical perspectives on politics. Lijphart's (1975) work on democracy in divided societies is an early example; more recent ones range from Scott's (1998) rumination on the organizing vision of states to the efforts of Swenson (2002) and Mares (2003) to explain how business groups approach social policy. Such works draw their insights from comparison across time and space but rarely rely heavily on the conventional comparative method. Instead, their methods are closer to what Verba (1967) called "disciplined," configurative inquiry - efforts to explore the similarities and differences in political phenomena across nations with a view to identifying new concepts, typologies or causal processes whose import has not hitherto been delineated. Moore's (1966) classic study of the origins of dictatorship and democracy is a good example. With some crudeness, one can reduce his analysis to a comparative study focused on a few key variables, but the principal contribution of the book was to illuminate the types of long-term historical processes that sow the seeds for regime change. Comparative politics has made its mark partly by painting new portraits of the world.

The second move is a shift toward alternative conceptions of what can be explained and in what causal explanation consists. Influenced by logical positivism, conventional presentations of the comparative method assumed that a good explanation was one that identified a few causal variables, operating independently of most others, with consistent effects in all cases and, if possible, attached to precise parameter estimates reflecting the relative influence of each. Occam's razor was the weapon of choice and parsimonious explanation highly valued (cf. Shively 1974).

> "... few scholars of comparative politics have ever relied heavily on the comparative method, at least as conventionally defined."

This approach to explanation was well-suited to the modernization perspectives of the 1960s and 1970s that located the critical determinants of many political outcomes, ranging from the stability of democracy to political attitudes, in socioeconomic developments. Readily measurable social or economic variables were expected to have powerful effects across contexts. In recent decades, however, comparative politics has embraced theories that associate political outcomes with more complex causal processes, such as sequences of strategic interactions or the feedback effects of path dependence (Pierson 2004). Compare recent accounts of the transition to democracy as a strategic interaction between reformers and established elites to the hydraulic theories of democratization of previous eras (Lipset 1959; Colomer 2000).

Partly as a result, prevailing views about what it means to give a good explanation of political outcomes are shifting. Influenced by critical realism, many scholars have moved away from the search for a few key causal variables toward the view that explanation consists in identifying the causal processes that lie behind political outcomes (cf. Yee 1996). Some see this as a problem of modeling complex interactions in which the sequencing of events matters and context effects condition action (cf. Scharpf 1997; Sewell 1996). Others advocate a search for the "mechanisms" that feed into causal processes, conceived as recurrent dynamics with social or psychological roots (McAdam et al. 2001). On the premise that different combinations of variables can produce similar outcomes, others move from variables-oriented explanations toward "case-oriented" explanations aimed at identifying such combinations (Ragin 1987; 2000).

In each of these inquiries, comparison remains central, but the object of inquiry has shifted toward the identification of a range of causal patterns that the conventional comparative method does not anticipate. Although debate about what kind of causal patterns political science should look for is continuing, the achievements of the field cannot be appreciated without acknowledging that it comprehends a wide range of views about what can be explained and what a good explanation embodies.

"... contemporary scholarship has moved beyond in two directions, one oriented to securing and analyzing more cases, the other focused on using more information from each case."

With regard to testing causal inferences, although some studies still exploit the comparative method to good effect, contemporary scholarship has moved beyond it in two directions, one oriented to securing and analyzing more cases, the other focused on using more information from each case.

Many comparativists have followed Lijphart's advice to make increasing use of statistical methods. Two developments have encouraged this step. One is the growth in cross-national data sets, fueled by interest in comparative political economy, political institutions, and democratization. The other has been the refinement and growing popularity of pooled-cross-sectional time-series analysis. Although it has more limitations than are sometimes acknowledged, by exploiting variation across time as well as space, such techniques allow scholars to apply regression analysis to propositions about nations even when the number of nations available for examination is small (cf. Beck and Katz 1995; Wallerstein 2000). As a result, the study of comparative politics is a much more statistical enterprise than it was in 1970.

Other scholars are using the innovative techniques Ragin (1987; 2000) has devised to apply the logic of the comparative method to a wider range of variables and cases than might otherwise be possible. Using Boolean algebra rather than the partial correlations of regression analysis, his methods yield different kinds of results, but ones that can be highly revealing about the necessary and sufficient conditions for cross-national political phenomena.

If such techniques expand the number of cases examined, the other step long taken by many scholars to bolster their causal claims entails more intensive analysis of each case. Here, the crucial move has been to go beyond the assumption, implicit in most formulations of the comparative method, that the most important information to be secured from a set of cases is about the correspondence between the values of a few key explanatory variables and the value of the dependent variable or outcome in those cases. Campbell (1975) paved the way many years ago, by noting that each case (defined as a unit of analysis displaying one value on the dependent variable) can yield many types of observations pertinent to assessing the validity of a causal theory.

There are interesting divergences of view about how to use the intensive examination of cases to improve causal inference. McKeown (1999) argues that intensive case studies should move beyond the "statistical worldview". Bates et al (1998) advo-

(Continued on page 4)

(Continued from page 3)

cate iterative procedures in which a theory is reassessed and revised as new pieces of information about the case are uncovered (cf. Elster 2000). Mahoney (2000) outlines a technique of "causal narrative" in which the adequacy of a causal theory is tested by asking whether the causal process it describes fits the unfolding sequence of events observable in successive cases. Pioneers in this area, Bennett and George (1997; George and McKeown et al 1985) propose a similar method of "process verification."

My own preference is for a relatively positivistic approach in which two or more causal theories are compared to one another, predictions are derived from each theory about key facets of the causal chain leading up to the outcomes, including the sequence in which events are likely to occur and the positions actors are likely to take, and observations are drawn from the cases and inspected to see how well they conform to the predictions of each theory (Hall 2003). We need more debate about precisely how process analysis should be done.

However, the key point is that many of the most fruitful studies in comparative politics employ some version of it, often in combination with the comparative method. The comparison of a small number of cases is and should remain central to work in comparative politics. But the trust placed in such studies often rests as much on the effectiveness of their process tracing, whether described as such or not, as on the correlations they find between the ultimate independent and dependent variables.

> "... my own preference is for a relatively positivistic approach in which two or more causal theories are compared to one another...."

It is striking, however, how rarely this dimension of comparative inquiry is discussed. Mesmerized by the comparative method and its statistical analogues, the field has been slow to acknowledge how important some version of process tracing is to most comparisons of a small number of cases. Fortunately, such issues are receiving more attention in the wake of the debates inspired by the efforts of King et al. (1994) to specify a unified approach to research design and the interest aroused by more complex theoretical accounts of the causal processes behind politics (cf. Brady and Collier 2004).

Today, the field is better defined not as one that makes preeminent use of a comparative method but as one that utilizes many methods to compare politics across nations. However, the attempts made to identify a distinctive comparative method have been valuable. Like the grit that sets an oyster in motion, they inspired a series of interchanges that have turned the field of comparative politics into one of the most lively sites for methodological debate in the discipline. That, in turn, has encouraged reflection about how conclusions are to be drawn about the political world, which ultimately strengthens the discipline.

The Proliferation of Comparative Survey Research

Introduction

One of the most important trends in comparative politics is the proliferation of multi-country survey research programs. This development has prompted some to question whether it is reasonable to expect meaningful and comparable results from large, multi-country surveys. In this symposium five prominent scholars take stock of this research, praise its potential, highlight its limitations, warn of its pitfalls, and offer constructive suggestions for executing it better.

Comparative survey research recovered slowly from the criticisms leveled against *The Civic Culture* (1963). For a decade, most surveys outside the U.S. were single-country studies rather than cross-national comparisons. Only with the appearance of the first Eurobarometer Survey in 1973, the European Values Surveys (1981), and the more ambitious World Values Surveys (1990-91; 1995-96; and 1999-2001) did this trend reverse. Inspired by the original European model, and with the steady expansion of the universe of new democracies around the world, there are now regional public-opinion barometers for Eastern Europe and the former Soviet Union, Africa, Asia, and Latin America. Moreover, in addition to these regional barometers, there has been a proliferation of cross-regional opinion surveys, including the Comparative Study of Electoral Systems Project, the Pew Global Attitudes Project, many USIA polls, and various NSF-funded projects.

In the symposium's first essay, Pippa Norris provides a comprehensive and insightful overview of the many political and intellectual factors that have contributed to the recent proliferation of cross-national survey research. Norris is ultimately optimistic that the proliferation of cross-national survey research will continue into the indefinite future and that, over time, the institutional survey infrastructure in developing countries will improve.

In the second contribution, Mitchell Seligson takes exception to the widely held view that comparative survey research languished from the mid-1960s to the mid-1970s and identifies four problems common to this research. One of Seligson's major criticisms of contemporary surveys is that by inferring individual-level relationships from aggregated survey results researchers fall victim to the ecological fallacy.

Buttressed with extensive data from the World Values Survey, Ronald Inglehart and Christian Welzel's essay implicitly responds to Seligson's criticism with a vigorous defense of the value of exploring relationships between survey aggregates and societal-level variables. The authors also return fire by warning against the individualistic fallacy.

In our fourth essay, James L. Gibson draws upon his background of conducting public opinion surveys in Russia and South Africa to caution about the difficulties of measuring concepts well, obtaining representative samples, reporting technical information and managing costs. Despite these significant difficulties, Gibson sees no viable alternative to survey research for those who wish to draw systemic inferences and conclusions about the similarities and differences of mass politics across countries.

In the final contribution, Debra Javeline, who has extensive experience carrying out survey research in the former Soviet Union and Eastern Europe, offers practical guidance for occasional practitioners, seasoned researchers, and consumers of public opinion survey research. Emphasizing the large body of expert knowledge that survey researchers have developed, she argues that those who intend to work with surveys must either acquire this expertise themselves or collaborate with those who are already expert.

Taken together, these reports from the front lines of comparative survey research urge comparative political scientists as well as practitioners to conduct and consume comparative opinion surveys cautiously. All our symposium authors agree that cross-national and cross-regional survey research has made and will continue to make important contributions to comparative politics scholarship. However, they also remind us that some common items do not measure their intended concepts well, some questionnaires are prone to response bias, some unscrupulous survey research firms have falsified responses, and some samples are not nationally representative. Consumers must learn to evaluate the quality of surveys, and practitioners and publishers must develop the habit of providing all the technical information that is needed to make such judgments. This symposium provides much-needed guidance that, if heeded, will enhance the contributions of cross-national opinion surveys to comparative politics.

From the Civic Culture to the Afrobarometer

Pippa Norris
Harvard University
Pippa_Norris@Harvard.edu

Survey resources for the systematic cross-national comparison of public opinion have expanded dramatically during recent decades. Many political and intellectual factors have contributed to this development. The expansion of the European Union to 25 members has played a direct role. The growth of electoral democracies has facilitated the study of public opinion and also the demand for commercial market research companies and non-profit social science institutes that are free from political interference and overt state censorship. International development agencies, such as the UNDP, the World Bank, and Transparency International, have increasingly recognized that programs seeking to expand democracy and good governance need to monitor public opinion as well as the standard "objective" developmental indicators.

Leaders in the field have had a decisive impact. Although it is invidious to single out a few, both Ronald Inglehart and Roger Jowell have played seminal roles, through generously initiating, managing, and sustaining major cross-national surveys which have had multiplier effects through funding public opinion institutes and training the next generation of fieldwork staff and survey analysts. Intellectual fashions have also contributed by eroding interest in the more old-fashioned historical-institutional tradition of area studies within specific countries and creating demand among the younger generation of researchers in Asia, Latin America, and Eastern Europe for more systematic cross-national comparison of political culture and behavior, conducted within varying institutional contexts.

"The growth of electoral democracies has facilitated the study of public opinion."

Perhaps the most recent spur has been the events of 9/11 and their aftermath, renewing American curiosity about public opinion in the rest of the world, particularly in regions like the Middle East where cross-national survey research has been scarce. These developments have gradually transformed the geographic scope of coverage, allowing comparativists to move "from nations to categories," one of the key but elusive goals of the sub-discipline. Nevertheless, the substantial advances in collaborative data collection have, somewhat curiously, moved ahead of the profession's exploitation of these resources.

Comparing Cross-National Surveys

Listed in Table 1 are eight survey series with significant potential for widespread use in comparative politics. We first compare basic information about these cross-national survey series and then briefly summarize their merits and shortcomings. The cross-national public opinion surveys included in Table 1 satisfy four criteria:

Sample: surveys are designed to monitor public opinion among a representative sample of the adult population in each country, excluding surveys of elites such as Transparency International's Corruption Perception Index,[1] or the CIVED civic education study among students;[2]

Coverage: surveys cover a substantial number of countries, ranging from 8 nations in the Asian Barometer to almost 80 in the World Values Survey;

Subject: the core questionnaire or battery of items focus on measuring political and social attitudes, values, and behavior, rather than other topics such as population or health surveys; and

Time-series: they constitute a recognizable series conducted over successive waves, rather than a single snapshot survey.

1. Eurobarometer

The Eurobarometers are the longest and one of the best-known series, constituting public opinion surveys conducted on behalf of the European Commission at least twice a year (spring and autumn) in all member states of the European Union since the early seventies. The geographic coverage has expanded with EU membership, monitoring social and political attitudes in the European publics. The Eurobarometer program was later enlarged by a small-scale Flash Eurobarometer and a Central and Eastern Eurobarometer; later replaced by the Candidate Countries Eurobarometer.

Eurobarometer raw data and documentation (questionnaires, codebooks etc.) are stored at the ICPSR and at the Zentralarchiv at Cologne and are available for social science research purposes (i.e. secondary analysis of the raw data) by the Social Science Data Archives. Survey results are regularly published by the Eurobarometer unit of the European Commission.

Bottom line: The Eurobarometer series is familiar to students of the politics of the European Union, but the full potential of this series is surprisingly under-utilized outside of this subfield, perhaps because of the complexities of handling over 60 large-scale surveys. The best exemplification of its potential use is the *Beliefs in Government* book series published in 1995 by Oxford University Press (Klingemann and Fuchs 1995, Niedermayer and Sinnott 1995, Borre and Scarbrough 1995, Van Deth and Scarbrough 1995, Kaase and Newton 1995).

2. World Values Survey/European Values Survey

The *grand dame* of cross-national surveys, the study is now entering its fifth wave. The World Values Survey is a worldwide investigation of sociocultural and political change. It has carried out representative national surveys of the basic values and beliefs of publics in almost 80 societies on all six inhabited continents, containing almost 80 percent of the world's population. It builds on the European Values Surveys, first carried out in 1981. A second wave of surveys, designed for global use, was completed in 1990-1991, a third wave was carried out in 1995-1996, and a fourth wave took place in 1999-2001. This study has informed more than 300 publications in 16 languages.

Bottom line: Still the only cross-national public opinion survey comparing all regions, with growing geographic coverage of the Middle East, Asia, and Africa. Time-series analysis is complicated by the fact that country coverage and some items vary across successive waves, and the 1981-83 first wave focused mainly on affluent nations. Nevertheless the WVS provides a benchmark for many developing societies, where for years it was the only widely available cross-national survey monitoring a wide range of social and political values. The new *Human Values and Beliefs* sourcebook (Siglo XXI Editores, 2004) makes the descriptive data accessible for the electronically challenged. Some of the most recent book publications from this include *Rising Tide: Gender Equality and Cultural Change Worldwide* (Inglehart and Norris, 2003) and *Sacred and Secular: Religion and Politics Worldwide* (Norris and Inglehart 2004), both with Cambridge University Press.

3. The International Social Survey Program

The International Social Survey Program (ISSP) is a continuing, annual program of cross-national collaboration. It brings together pre-existing social science projects such as the General Social Survey at NORC and the British Social Attitudes series and coordinates research goals among the consortium, thereby adding a cross-national perspective to the individual, national studies. Since 1983, the ISSP has grown to 38 nations, including many industrial and post-industrial societies. The focus is the inclusion of a thematic annual module with a battery of items carried in existing social national surveys, with the annual theme covering rotating issues in the social sciences, such as national identity, the role of government, religion, the environment, work orientations, and gender roles. Considerable attention is paid to standardizing the social and demographic background information in the surveys.

Bottom line: The ISSP has a far more limited geographic range than the WVS, and a narrower thematic focus than the WVS or the EB, but nevertheless, if the thematic topic fits your research needs, the survey provides considerable depth on each topic. Illustrative books include Roger Jowell, Lindsay Brook, and Lizanne Dowds, eds., *International Social Attitudes: The 10th BSA Report* (Dartmouth, 1993).

> "The multiplicity of surveys is to be welcomed by facilitating replication both across years and across nations."

4. The Comparative Study of Electoral Systems

The Comparative Study of Electoral Systems (CSES) brings together an international team of collaborators who have incorporated a special battery of survey questions into the national election studies. Data from each of the separate election studies is coordinated, integrated and cleaned by the Center for Political Studies, Institute for Social Research, at the University of Michigan. The dataset is designed to facilitate the comparison of macro and micro-level electoral data. Module 1 of the CSES (1996-2001) allows comparison of a representative cross-section of the electorate in 37 legislative and presidential national elections in 32 countries. The geographic coverage is remarkably diverse, ranging from Belarus and Ukraine to Canada, Australia, and Belgium. The focus on voters' choices, the cross-national integration, and above all the timing of the data collection (within a year following each of the elections), provide a unique opportunity to compare voting behavior in a way that is not possible through other common sources of comparative data such as the World Values Survey. Fieldwork, data collection, and integration of the second module (2001-2005) is in process, with some country datasets already released, and planning for Module 3 is currently under way.

Bottom line: The CSES is the best data currently available for cross-national electoral analysis but analysis is complicated by the diverse range of regions, political systems, and levels of democracy included in the study. This requires a "most different" comparative strategy, rather than the familiar regional/area approach. The integration of the data, for example the demographic and social coding, is also far more complicated than in a single-funded survey such as the Eurobarometer. The main strength is the capacity for multi-level analysis combining analysis of political behavior within contrasting institutional contexts.

5. The Globalbarometers

Rather than a single entity, these are actually four separate regional series, loosely coordinated, and originally inspired by the Eurobarometer model. The *New Europe* series coordinated by Richard Rose at Strathclyde University has focused upon monitoring the process of cultural change in political and economic attitudes following the breakdown of communism. The survey is conducted every year in selected Central and Eastern European countries and it has resulted in numerous papers and books published by Professor Rose and his collaborators.

Under Marta Lagos (MORI, Santiago), the *Latin Barometer* has conducted pioneering work monitoring annual trends in attitudes towards democracy. The series started with 8 nations in 1995 and subsequently expanded in 17 countries in the region. Although an invaluable resource, and founded as a non-profit institution, the electronic data is less easily available for secondary analysis than other surveys compared here, and as a result it has had less impact on Latin Americanists than might be expected given the topic and the quality of the data.

The Afrobarometer is one of the newer members of the stable, pioneered by Michael Bratton, Robert Mattes, and E. Gyimah-Boadi, who developed networks of surveyors in many countries that had never had any previous social scientific surveys of political and social attitudes. The Afrobarometer also serves as a model of transparency for others in releasing full information about the work in progress, including questionnaires, publications, funding, and associates, as well as depositing all data through archives and its own dedicated website.

Lastly, the *Asian Barometer* is the latest to join the network, sharing similar concerns to monitor public attitudes towards democratization and economic development, with 8 nations coordinated in the survey by Yun-han Chu in Taiwan.

Bottom line: At last, survey research is covering large parts of the developing world, such as Africa, which were previously neglected, and thereby building up the infrastructure of experienced fieldwork teams and survey analysts that will pay dividends in future. The surveys facilitate cross-national comparisons among a group of collaborators while also retaining regional autonomy to focus on specific themes of interest to each area. One of the most important remaining challenges is to make sure that this data is available and utilized by the social science communities within each region, by equipping the next generation of graduate students with the necessary intellectual frameworks, skills, and infrastructure to exploit the data. The first book emerging from the Afrobarometer series is Michael Bratton, Robert Mattes and E. Gyimah-Boadi *Learning about Reform in Africa: Public Opinion, Democracy, and Markets* (Cambridge University Press 2004).

6.The European Social Survey

The European Social Survey (the ESS) is a new, academically-driven social survey designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behavior patterns of its diverse populations. The survey covers at least 21 nations in Western and Central Europe and it will employ the most rigorous methodologies. The survey contains a core that will be replicated every two years plus rotating thematic modules, allocated to teams of scholars on a competitive basis. The organization of the survey is a model of transparency as well as employing the highest standards in sampling and fieldwork practices, and carefully standardizing the collection of social and demographic background data.

Bottom line: In time, this will come to be regarded as the Rolls-Royce of cross-national surveys. The central coordination and funding of the ESS, the care in crafting and testing the questionnaire, and the development of additional contextual data, provides a model that other cross-national survey researchers will seek to emulate (if only they had the resources). It is a pure pleasure for any analyst to work with this survey.

7. The Pew Global Attitudes Survey

In response to the aftermath of 9/11, attention in the United States has turned increasingly towards understanding how the world (particularly the Muslim world) views America. The “clash of civilizations” thesis has also spurred greater interest among the international relations community into issues of global cultural similarity and difference. Andrew Kohut at the Pew Center for the People & the Press responded to this with the Pew Global Attitudes Survey, an annual attempt to monitor public opinion in many countries, using commercial market research companies. The *Pew Global Attitudes Project* is a series of worldwide public opinion surveys, originally of more than 38,000 people in 44 countries in 2002, and expanded in 2003 with additional surveys to a total of nearly 75,000 people among the 50 populations surveyed (49 countries plus the Palestinian Authority). The project encompasses a broad array of subjects ranging from people’s assessments of their own lives to their views about the current state of the world and important issues of the day.

Bottom line: The published reports are available online, attracting considerable media attention and interest in the State Department and the policy community, but the data is not yet released for secondary analysis.

8. Gallup International *Voice of the People*

The last survey, coordinated by Gallup International, is similar in orientation to the Pew survey but with a commercial orientation. In 2002 Gallup International conducted a worldwide survey of 60 nations monitoring attitudes towards issues such as the environment, terrorism, global issues, governance, and democracy. In 2003 this survey was conducted again covering Western Europe, the USA, and Canada but also Africa, the Middle East, Asia, Eastern Europe, and Latin America. Highlights of the results are published on their website but the published report (containing detailed cross-tabulations) and the electronic data are available only for purchase and Gallup International offers the opportunity for clients to add items to the questionnaire, also at a cost. It is not possible to evaluate the quality of the detailed methodology, sampling, and fieldwork practices in countries where surveys are uncommon, such as in the Middle East and Africa, from the publicly available information released on Gallup’s website.

Bottom line: Both Pew and Gallup are breaking new ground by expanding their geographic coverage in ambitious attempts to monitor public opinion around the world. This is a welcome addition to the available resources but it remains too early to evaluate the quality of these surveys, and unfortunately access to the data remains limited.

Conclusions

Opportunities for cross-national survey research have been transformed over the last twenty years. The multiplicity of surveys is to be welcomed by facilitating replication both across years and among nations. Some of the more commercial initiatives may fail, for example, if America withdraws into itself and turns away from the world again, in its periodic cyclical fashion. Yet, it seems likely that the underlying momentum will continue in subsequent decades, as younger generations of social scientists trained in survey methods and public opinion analysis will develop in each region.

Obvious questions can be raised about the quality of sampling and fieldwork in poorer, developing nations which have not built up experienced market research companies and social science institutes. But the replication allows some cross-checks to be developed across different surveys. Multiple familiar questions can be raised about the quality of questionnaire translations and the employment of equivalent standards across different nations - debates which have been with us ever since *The Civic*

“... it seems likely that the momentum will continue in subsequent decades, as younger generations of social scientists trained in survey methods and public opinion analysis will develop in each region.”

Table 1: Key features of the cross-national series of surveys

Series	Series started (i)	Frequency	Nations (latest survey) (ii)	Data downloadable (iii)	Coordinating Organization	Online resources Http://
Eurobarometer	1973	Bi-annual	25	Public archives	Antonis PAPACOSTAS Directorate General Press & Comms, European Commission	Organizing & reports: http://europa.eu.int/comm/public_opinion/ Data and continuity guides from ZUMA, Cologne Archive: www.gesis.org/en/data_service/eurobarometer
World Values Study-European Values Study	1981-1983	Approx. 5 years	79	Public archives	Ronald INGLEHART, Institute of Social Research, University of Michigan	Organizing and data; www.worldvaluessurvey.org/
International Social Survey Program (ISSP)	1985	Annual	38	Public archives	Secretariat: Bjørn HENRICHSEN, Norwegian Social Science Data Services (NSD), Bergen	Organizing: www.issp.org/ Data and continuity guide from the ZUMA Cologne Archive: www.gesis.org/en/data_service/issp/
Comparative Study of Electoral Systems (CSES)	1996-2001	Module every 5 years	31	Public archives	Secretariat: David HOWELL, ISR, University of Michigan. Chair: Ian McAllister, ANU	Organizing and data: www.umich.edu/~cses/
Global-barometers, including:						www.globalbarometer.org/
New Europe Barometers	1991	Annual	16		Richard ROSE, CSPP, Strathclyde University	www.cspp.strath.ac.uk
Afrobarometer	1999	Annual	15	Public archives	Michael BRATTON (Michigan State), Robert MATTES (IDASA, SA) and Dr E. GYIMAH-BOADI (CDD Ghana)	www.afrobarometer.org
Latino-barometer	1995	Annual	17	Tables only	Marta LAGOS, MORI, Santiago	www.latinobarometro.org
Asian barometer	2001	Annual	8		Yun-han Chu, Taiwan	www.eastasiabarometer.org
European Social Survey (ESS)	2002	Biennial	21	Public archives	Roger JOWELL, Center for Comparative Social Surveys, City University	Organizing: naticent02.uuhost.uk.uu.net Data from the Norwegian archive: http://ess.nsd.uib.no .
Pew Global Attitudes Survey	2002	Annual	49	Only tables released	Andrew KOHUT, Director, The Pew Research Center for the People & the Press	people-press.org/pgap/
Gallup International *Voice of the People*	2002	Annual	60	Only tables released	Meril JAMES, Secretary General Gallup International	www.voice-of-the-people.net/

Notes: (i) In some cases there were often pilot studies and forerunners, such as the European Community Study, but this date is the recognizable start of the series in its present form. (ii) The number of countries in each survey often varies by year. (iii) If not deposited in public archives or directly downloadable, access may be available from the survey organizers on request, but there might be charges for access.

Culture. Yet in counterbalance there are certain distinct practical advantages associated with conducting surveys in developing nations, namely much lower refusal and non-response rates (currently approaching record levels for opinion polls conducted in the US), as well as relatively low budgets. Over time, as greater experience is gained, and as an institutional survey infrastructure is developed in the social sciences, these initiatives will mature.

Notes

[1] Although Transparency International have now launched the Transparency International Global Corruption Barometer among a representative sample of the public. This is intended to be an annual series monitoring public opinion on this topic, carried as part of Gallup International's *Voice of the People Survey* in 2002 and 2003. http://www.transparency.org/surveys/index.html#barometer

[2] Details about the International IEA CIVED: http://iea-dpc.de/Home_e/Studies/CIVED_e/cived_e.html, and http://nces.ed.gov/Survey/ CivEd/.

Comparative Survey Research: Is there a Problem?

Mitchell A. Seligson
Vanderbuilt University
seligsonm@yahoo.com

The editors of APSA-CP made two assertions in their charge to the participants in this symposium. First, they made the claim that "comparative survey research recovered slowly from the criticisms leveled against *The Civic Culture* (1963)." Second, they stated that "for a decade, most surveys outside the U.S. were single-country studies rather than cross-national comparisons." In this essay I take issue with both of these assertions, and then go on to raise several cautionary flags about the ways in which recent surveys on democracy have been carried out and interpreted that speak directly to the central question that the editors ask: "Is it reasonable to expect meaningful and comparable results from large, multi-country surveys?"

The Past

The first claim made by the editors is that the flaws in the classic work of Almond and Verba were so serious that they paralyzed, or at least largely stifled, comparative survey research for a decade. In fact, if imitation is the highest form of flattery, the record shows just the opposite. Prior to the publication of *The Civic Culture,* social scientists were burdened with "national character" as the only concept in their bag of tools that they could use to explain variations in political beliefs across nations. That research, as we know, left us with little more than nuggets, such as "Americans are industrious," "Germans are obsessive about details" and "the French," of course, "are lovers."

Then along came the seminal work of Almond and Verba and the term "political culture" became a household word in political science, not merely replacing "national character," but introducing it as a systematic and, most importantly, measurable variable in our efforts to explain similarities and differences among nations. As a result, there emerged a veritable explosion in political culture studies. When one examines the titles of early political culture dissertations, one finds that initial evidence appears to support the contention that comparative survey research died with the publication of *The Civic Culture*. Indeed, many, if not most of these dissertations refer to a single country

"... many of the nations we set out to study back then were in the grip of dictatorships, which made carrying out national-level surveys difficult, if not downright impossible...."

in their titles. However, closer examination of the dissertations themselves, together with articles published in scholarly journals of the day, reveals two things. First, a number of these studies were based on the Almond and Verba data set itself, but often the analysis was of only one of the five countries. Second, many other studies collected new datasets in single countries, but used the very questions (or versions of those questions) that Almond and Verba had asked, and in a number of those studies results were explicitly compared to those of Almond and Verba.

It is true that very few (perhaps no) multi-nation studies were carried out in the years immediately following the publication of the Almond and Verba classic, but this was not because of the criticisms leveled against it. Rather, four other factors were responsible. First, the mere presence and availability of the five-country data set provided the raw material for scores of researchers who analyzed and reanalyzed that material in the decades that followed its public release. Second, comparative politics graduate students in those days were required, by the norms of the day governing future employment prospects, to become area studies experts, and were expected to produce a dissertation that was a monographic work on a single country that they could claim to know well. Third, those who aspired to carry out multi-nation studies faced the hard realities of limited funding. The foundations that supported the original study moved on to support the new "next big thing," while the NSF political science program opened its doors only in 1965.

A fourth factor was perhaps the single most important one: many of the nations we set out to study back then were in the grip of dictatorships, which made carrying out national-level surveys difficult, if not downright impossible, and certainly dangerous. In the mid-1970s, when John Booth and I set out to do comparative work on political participation in Latin America, we searched the field for scholars who had conducted surveys in the region. What we found, and later reported in two edited volumes, was that dozens of scholars were in fact doing surveys, many of them including questions drawn right out of *The Civic Culture*, but that they were limited by the constraints of funding and, more importantly, by fears of incurring the wrath of military dictators. There was a series of small-scale, but nonetheless comparative projects in which two villages or a few towns were compared to each other, and to percentages taken from the Almond and Verba study. These were, then, comparative studies, but not cross-nationally comparative in most cases.

The Present

Yes, *The Civic Culture* was a flawed work, but it was a pioneering study that is certainly among the most cited books in political science over the past forty years. Its flaws largely encouraged rather than discouraged research in a field that is today vibrant and growing. Yet, it is time for us to take stock about the limitations of many of the large-scale, multi-nation data sets that are today coming to dominate the field. In this concluding section I discuss the key problems that have emerged in these large multi-country data bases and suggest some solutions. Specifically, I examine four issues: 1) problems in measuring support for democracy, 2) quality of data, 3) transparency in the availability of data sets, and 4) problems of aggregation.

Measuring support for democracy

Years ago, when social psychologist David McClelland (McClelland1976) developed the notion of the "achievement motive" to explain variations in national growth, he warned us that to understand what drives people one has to use unobtrusive measures. If that warning has any relevance it is that if one is trying to understand support for democracy, one ought not to use the term "democracy" in one's surveys. Almond and Verba avoided the use of the term "democracy" entirely in their questionnaire, and instead focused on values and behaviors that they believed explain how a democratic way of life emerges and is sustained. Today, however, it has become de rigueur to ask questions-about whether "democracy is the best system," or "democracy is preferable to dictatorship," or "how well does democracy function in this country?" and similar items.

We face two problems when we ask respondents directly if they support democracy. First, there is the overwhelming problem of "social desirability response set." Many respondents seek to give what they perceive as the "right" answer to the interviewer. If democracy is thought to be "good" an important portion of pro-democracy response could be an artifact of this form of response set, completely independent of respondents' actual preference for regime type. Second, there is the problem of respondent variation in the definition of "democracy." I recall talking to a Chilean Pinochet supporter who during the darkest days of that dictatorship forcefully told me that the Pinochet regime was the most democratic one that Chile had ever had, and indeed was the most democratic regime in all of Latin America. Years later, when democracy had been restored in Chile, I investigated definitional variation in a comparative study of Mexico, Chile and Costa Rica (Seligson 2001). When asking respondents to define democracy as a follow up question, I found, that whereas in Costa Rica, overwhelmingly "freedom" was given as the reply, in Chile and Mexico very different responses emerged, including, in the Chilean case, a strong identification of democracy with free trade.

The solution that I think works best for surveys that try to gauge popular support for democracy is to dispense entirely with using the term in the questions that are crafted, and instead, to focus on the operational definitions of a democratic political culture that our theories tell us we expect citizens to support. I have stressed, for example, the combination of political tolerance and a belief in the legitimacy of the institutions of democracy. Unless citizens are tolerant of minority rights, they will not be willing to allow minorities to compete for and actually win political power. I also argue that if citizens do not believe in the legitimacy of their key political institutions the stability of their political system is hard to ensure. My research in Latin America in recent years has found that only a small percentage of all those interviewed hold the dual combination of tolerant attitudes and a belief in the legitimacy of their political system, Costa Rica being an important exception. Far greater numbers are both intolerant and unsupportive of the system (Seligson 2000).

Data quality issues

All survey data sets contain error. These errors occur because of flawed sample frames, poor question design, poor interview techniques, flawed coding of responses, and errors in data entry. It should not surprise us, then, that our explained variance is often very low; even with the "perfect" theory, random error in data attenuates significant correlations. No survey will ever fully overcome all of these problems, but rigorous quality control can minimize them.

Unfortunately, in recent years, with the explosion of surveys in the developing/democratizing world, quality control has suffered, in some cases precipitously so. When the "noise" is greater than the "signal," our data become uninterpretable. In effect, the "signal-to-noise ratio" is so poor that we cannot, in effect, any longer hear the "music." One illustration of this problem will be illuminating. A basic rule in the survey field is that to minimize errors in data entry, all data sets must be verified (i.e., double entry by two different data entry clerks). All data-entry software packages of which I am aware make provision for data verification and produce automatic alerts when the first entry is inconsistent with the second. All of us who commission surveys in the developing world, I presume, require (and pay for) data verification. Yet, it has been my experience in Latin America that verification is the exception, even when we pay for verification. In order to attenuate this problem, I use data entry packages - like the U.S Census Bureau's CSPro 4.0 - that produce codes that reveal, case-by-case, whether the data set has been verified. Second, I conduct an audit of the data entry itself by randomly selecting some 100 interviews and comparing them with the actual questionnaires. I compare the written replies on the questionnaire with their assigned codes, and then compare them to the datasets themselves. This tedious work quickly reveals whether the coding and data entry have been done well and often reveals instances of falsified data.

> “Scientific advances are accepted only if they can be replicated. [...] In the field of multi-national survey research, we increasingly see a movement in the opposite direction, with datasets being embargoed ...for years.”

Transparency

Scientific advances are accepted only if they can be replicated. We all recall the rapidity with which our physicist friends debunked "cold fusion" when they could not replicate the results of the original investigators. In the field of multi-national survey research, we increasingly see a movement in the opposite direction, with datasets being embargoed, sometimes for years. The pressure to publish is, of course, a reason why such embargoes exist, and researchers who have expended great effort to obtain the funding to gather their data understandably do not want to just give it away. In addition, draconian regulations of some of the Institutional Review Boards to which we are all subject, sometimes place insurmountable barriers in the way of data sharing. Yet, if our enterprise of advancing our science is to succeed, we must find ways of overcoming the privatization of data, especially that which is paid for with public funds. The recent release of a CD version of the World Values Survey is exemplary (Inglehart, et al. 2004) in this respect.

The levels of analysis problem

International relations experts have long been concerned with what they have called "the levels of analysis problem," but those of us who work in comparative politics have been far less sensitive to this issue, probably because so many of us work on individual countries. The widespread availability of multi-nation data sets, however, has led to confusion as to the most appropriate way to treat such data sets. Ron Inglehart and I have recently debated this issue (Seligson 2002; Inglehart and Welzel 2003) and there is no need to review those arguments here other than to reemphasize two points. First, when we aggregate attitude items into scales based on national means, without first determining a positive and significant correlation of individual components with one another, we risk creating statistical artifacts. Second, if we are to demonstrate to our new institutionalist colleagues that values do have an independent role in explaining democratic outcomes, and are not merely a byproduct of the creation of democratic institutions, we need to use measures of attitudes from the period prior to the dependent variable (e.g., level of democracy) that we are attempting to explain.

The central emerging levels-of-analysis issue that confronts us today, however, is how to treat individual-level data in multi-nation regression analyses. Individual observations at the cross-national level are clearly not independent of one another (respondents are residents of nations with particular levels of wealth, income distribution, and welfare, as well as unique combinations of political institutions and historical background). Unless we successfully deal with this issue in our regressions, we are likely to find significant associations that are merely artifacts of our multi-level designs. Two approaches have emerged in recent political science journal articles, one of which uses multi-level statistical routines, such as the increasingly popular HLM (Raudenbush, et al. 2000). However, the stiff requirements for the number of units at each level prevents those of us who work with a small number of countries to meet those requirements. This has pushed some scholars to move in a second direction, namely that of using the STATA routines that treat the nation as a single cluster in a multi-stage stratified sample. Doing so can help account for the "intra-class correlation" that Kish (1965) helped us understand so many years ago in his pioneering work on sample design. It is not at all clear at this juncture (to me at least) which approach will turn out to be the more fruitful, given the known limitations of both. Comparative survey research is fortunate to have an unprecedented richness of databases, and we are likely to see rapid advances in our understanding of politics as a result. This essay has raised some cautionary flags that hopefully will help us move in that direction with more precision.

What Insights Can Multi-Country Surveys Provide about People and Societies?

Ronald Inglehart
University of Michigan
ringlehart@umich.edu

Christian Welzel
International
University Bremen
c.welzel@iu-bremen.de

Large-N cross-national surveys can provide insights into human behavior that are otherwise unobtainable. They make it possible to perform empirical analyses of cross-level linkages, testing such questions as: "Are specific individual-level beliefs conducive to the emergence and flourishing of democratic institutions?" "Is economic development linked with declining religious beliefs and more secular worldviews?" "Are certain individual-level beliefs conducive to economic growth?" These questions have been debated by students of political culture, secularization and modernization for decades. Since they involve individual-level beliefs, any conclusive answer requires survey data. Because they also involve system-level variables such as democratic institutions or levels of economic development, they require data from a broad range of societies and, insofar as they involve changes over time, they also require data from multiple waves of surveys.

Only recently have cross-national survey projects such as the various Barometer surveys now being carried out in Europe, Asia, Latin America and Africa; the ISSP surveys; the Comparative Election Surveys; and the World Values Survey (which has just released an 80-nation dataset) begun to provide the data needed to answer such questions. The availability of comparable, cross-national survey data from countries covering the full range of economic, political and cultural variation will not end these debates, but it will make it possible to move away from relying on stereotypes and guesswork, allowing us to base our arguments on replicable evidence.

Needless to say, multi-country surveys must be designed and conducted carefully in order to produce useful results. No research method is exempt from these requirements, but the opportunities for error are multiplied when one moves from carrying out research in a single society to doing it in many different societies. Probably the most obvious difficulty is the fact that certain words and key concepts have different meanings and connotations in different societies. Producing cross-culturally comparable questionnaires is a complicated and challenging process. So too are the problems of sampling and interviewing in widely different settings. There is a huge literature on this topic, which we will not summarize here. Having helped design and analyze literally hundreds of cross-national surveys, we conclude that these problems are real and very important, but they are not insurmountable. Consequently, we will focus on the question: What insights can be uniquely obtained from multi-country surveys, assuming they are designed and conducted properly?

We start from the self-evident but significant premise that human societies are composed of people. This means that societies are driven by patterns of mass behavior. This behavior is rooted in people's prevailing psychological orientations, including their beliefs, values and motivations. Hence, in order to understand how societies function and develop, one needs to understand how their traditions, institutions and regulations are anchored in people's beliefs, values and motivations. Survey research makes such understanding possible.

Some additional considerations are useful in exploiting the potential of cross-national surveys. First, it is necessary to deal with representative surveys that measure the motivational and behavioral patterns of entire countries, if countries are the unit of analysis. Many of the most interesting variables, from democratic institutions to economic growth rates, exist at the national level, which means that representative national surveys are needed in order to analyze population-system linkages.

Second, it is useful to focus on survey questions that tap into deeply seated values and beliefs, rather than opinions that fluctuate from one day to the next. Opinions are highly susceptible to the problem of "non-attitudes" (Converse, 1970), and this is particularly true of political ones. Specific political issues are usually remote from people's daily lives and tend to evoke superficial opinions. Moreover, salient political issues are generally linked to nation-specific and period-specific political agendas, which reduces their comparability over space and time. It is thus useful to design questions that tap into relatively deeply rooted values, such as people's beliefs about gender roles, religion, personal liberty, state authority, and peoples' trust and tolerance towards others. Such social values not only have a higher comparative value across space and time but, surprising as it may seem, they also tend to have more impact on important system-level variables than do political orientations. Social trust, for example, has a stronger impact on "good governance" than political trust, though the latter would seem to be more directly relevant (Newton, 2001). Similarly, attitudes such as tolerance, trust, and emphasis on self-expression are much stronger predictors of system-level democracy than is overt support for democracy itself (Inglehart 2003).

Cross-cultural surveys are indispensable to test and develop general theories of society. Economic theorists from Landes to Barro, social capital theorists from Putnam to Fukuyama, and democratic theorists from Lipset to Dahl argue that a society's enduring cultural traditions play an important role in shaping economic, social and political development. The only way to test such claims is to conduct cross-cultural surveys - thus making it possible to analyze culture as a variable. Similarly, cross-cultural surveys that are repeated on a regular basis are required to test general theories across space and time.

As academic disciplines mature, they diversify theoretically, thus giving rise to the need for increasingly encompassing theories that can integrate various subfields of the discipline. Recent integrative theories that touch on various debates in comparative politics include Tsebelis's veto-player theory in the field of governmental institutions and Esping-Andersen's theory of welfare state regimes in the field of public policies. A similar concentration is occurring in the field of political culture, where much of the current discussion focuses on the intersecting concepts of social capital, civil society, and civic culture. This overlapping area might be described as the area of "civicness" (Putnam, 1993), which concerns the psychological and normative orientations that enable people to be economically productive, socially cooperative and politically democratic. A widely-shared consensus is emerging that these three sorts of beneficial mass behavior are rooted in a specific bundle of "civic" orientations, such as social trust (Norris, 2003).

Growing evidence that the orientations of ordinary people are important for a society's level of economic productivity, gender equality, and democratic consolidation has increased the relevance of survey data. Accordingly, economists, sociologists, psychologists and political scientists are increasingly using data from the World Values Surveys and other large cross-national surveys to analyze social and political change. Questions of civicness have become a common forum that integrates various disciplines that share an interest in human progress and well-being. The insight that the orientations of ordinary citizens matter is central to the concept of "human development" (Sen, 1999), which integrates economic, social and political aspects of progress into a single people-centered concept of development (Welzel, Inglehart and Klingemann, 2003; Welzel, 2003). There is considerable debate about which civic orientations are most effective in leading people to behave in productive, cooperative and democratic ways, but the concept of civicness nevertheless serves as a common point of reference.

> "... it is extremely unlikely that one would find such strong linkages as those in Figure 1 if survey data were contaminated by massive measurement error."

In the context of survey research, the theoretically most far-reaching type of analysis is done in studies relating aggregate survey data to societal-level phenomena. Knack and Keefer (1997), for example, use aggregate measures of social trust to explain economic growth. Inglehart (1997) has explored the linkage between civic norms and the longevity of democratic institutions. Another example is shown in Figure 1, which demonstrates the relationship between individual-level self-expression values and "good governance" - a composite indicator of "voice and accountability" developed by the World Bank. This indicator summarizes various measures of the extent to which people can select, replace and monitor governments (Kaufman, Kraay and Mastruzzi, 2003).

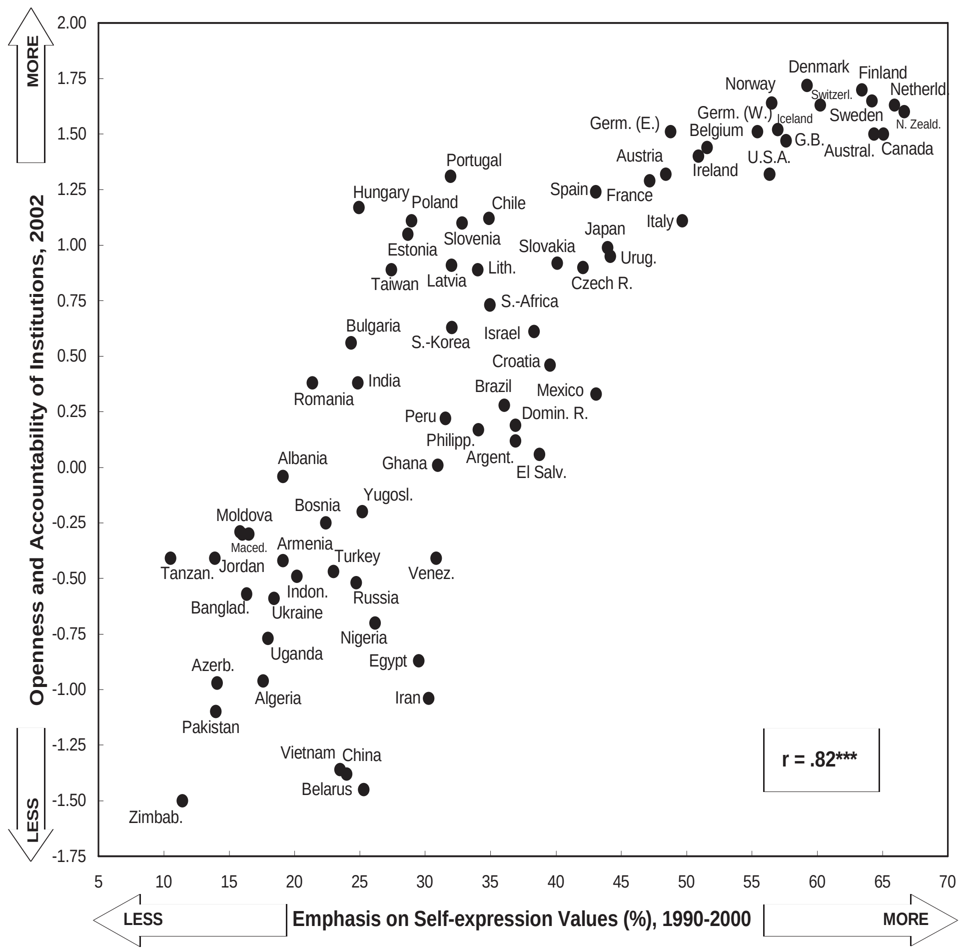

Figure 1

Our indicator of self-expression values is taken from the earliest available survey among the second to fourth wave of the World Values Survey (1990-2000). This measure reflects the percentages of people who: emphasize freedom and participation; tolerate sexual liberty; sign petitions; and trust other people and report high life satisfaction. As Figure 1 demonstrates, there is a strong linear relationship between mass emphasis on self-expression values within a society and the extent to which their respective political systems provide open and accountable institutions.

The relationship depicted in Figure 1 can be interpreted in various ways, but theoretical considerations suggest that the linkage between open and accountable institutions and mass-level self-expression values reflects the impact of values on institutions rather than the reverse. For, while self-expression values inherently lead people to demand open and accountable institutions, there is no reason why the presence of such institutions would tend to create these values. Moreover, in Figure 1, self-expression values are measured prior to open and accountable institutions. Thus, the relationship does not reflect the impact of institutions on values, unless both self-expression values and open and accountable institutions have a third common cause, such as economic development or prior democracy. However, even controlling for prior measures of democratic institutions and economic development, the effect of self-expression values remains significant: when one controls for per capita GDP in 1995 and the number of years spent under democracy, mass-level self-expression values still have a significantly positive effect on open and accountable institutions (partial r = .48, significant at the .001-level).

Regardless of whether mass-level self-expression values create or reflect open and accountable institutions, the fact that these two phenomena are so strongly related to each other is an important finding with two implications. The fact that a society's institutional performance is strongly linked with mass-level value orientations, as measured by completely independent methods, helps validate the survey evidence; despite the undeniable difficulties of cross-cultural survey research, it is extremely unlikely that one would find such strong linkages as those in Figure 1, if the survey data were contaminated by massive measurement error.

"... while survey researchers tend to be over-sensitive to the ecological fallacy, they tend to ignore the individualistic fallacy...."

Second, the finding that ordinary people's value orientations shape or reflect a society's basic institutional traits, points to the existence of a population-system linkage that ties genuine system characteristics, such as the quality of democratic institutions, to the central tendencies of given populations. Analyzing such population-system linkages is crucial to understanding how societies operate and develop.

The thesis that political institutions are rooted in ordinary people's value orientations is far from new; indeed, it is the central claim of civic culture theory. But this claim had been debated from the start because the empirical evidence was lacking. With an N of only five cases, it was impossible to make a conclusive test of Almond and Verba's claims. Evidence from recent cross-cultural surveys suggests that the central claim of the civic culture thesis is valid: certain cultural attitudes (though not precisely the ones that Almond and Verba emphasized most strongly) seem to play a crucial role in how well democracies function.

A major value of multi-country surveys is that they can demonstrate or disconfirm cross-level linkages that tie system-level characteristics to population tendencies. Such linkages are central to general social theories, illuminating the psychological dimension of social reality, which is not covered by economic or institutional data. But even today, the analysis of cross-level linkages remains poorly understood. Indeed, it is sometimes assumed that cross-level analysis often commits an "ecological fallacy." Thus, Seligson (2002) claims that the strong aggregate-level correlations that Inglehart found between political culture and stable democracy are "spurious" because he finds weak individual-level correlations between Inglehart's indicators of political culture and individual-level support for democracy. Seligson's analysis equates individual-level support for democracy with the presence of democratic institutions, a system-level phenomenon. Surprising as it may seem, however, individual-level lip service to democracy is only weakly linked with societal-level democracy. At this point in history, democracy has a positive image almost everywhere, but these favorable opinions are often superficial and, unless they are accompanied by more deeply rooted orientations of tolerance, trust, and a participatory outlook, the chances are poor that effective democracy will be present at the societal level.

Ironically, Seligson's conclusion exemplifies precisely the sort of cross-level fallacy that Robinson (1950) warned against. The central point of the ecological fallacy thesis is that strong aggregate-level relationships are not necessarily reproduced at the individual level. When Robinson was writing, districts with large percentages of African Americans (then located mainly in the South) generally elected segregationist candidates. But, as Robinson demonstrated, this relationship was not reproduced at the individual level: Blacks did not vote for segregationist candidates. This did not mean that the aggregate level relationship was "spurious"; no one questions the fact that districts with large numbers of African Americans really did elect the worst sort of segregationists, in a pattern of repression that endured for decades. The claim that an aggregate-level finding must be reproduced at the individual level is groundless, as Robinson demonstrated more than 50 years ago.

Interestingly, while survey researchers tend to be over-sensitive to the ecological fallacy, they tend to ignore the individualistic fallacy in assuming that any relationship found at the individual level will apply at the societal level. This is demonstrably wrong (Inglehart and Welzel, 2003). For example, much of the research on support for democracy assumes that individual-level lip service to democracy provides a reliable measure of how securely established democratic institutions are at the societal level. At first glance, this seems very plausible - but the relationship between mass attitudes and stable democracy can only be tested at the societal level. Moreover, it turns out that there is a surprisingly weak linkage between overt support for democracy, and actual democracy as measured by the Freedom House scores or other indicators of the degree and stability of democratic institutions. Empirically, the Albanians and Azerbaijanis are more likely to say favorable things about democracy than the Swedes or Swiss. This does not mean that democracy is more securely established in Albania than in Sweden - it reflects the emergence of so-called "critical citizens" (Norris [ed.], 1999) among stable democracies, coupled with the fact that, at this point in history, saying favorable things about democracy has become the socially desirable response in most societies.

In short, relationships that apply at the individual-level do not necessarily apply at the societal level, and vice versa. Properly aggregated survey data from a large number of societies are needed to analyze the impact of mass beliefs and values on democratic institutions or any other societal-level phenomenon. The analysis of such linkages is central to testing key hypotheses in democratic theory. Moreover, the availability of data from large-N cross-national surveys makes it possible to analyze cross-level linkages more effectively today than ever before.

Ruminations on Survey Research in Comparative Politics: Difficulties and Desiderata[1]

James L.Gibson
Washington University in Saint. Louis
jgibson@wustl.edu

I have been asked by the editors to address the question of whether extant cross-national survey research data bases are capable of producing meaningful and comparable results. This query is entirely appropriate, although it pertains with equal force to *all* cross-national data, not just survey data. So far as I can see, there is nothing about survey data rendering them uniquely suspect or problematical as compared to data on government spending, the attributes of institutions, or the policy stances of political parties.

I should note initially that the data bases we are discussing have been created by different organizations and for a variety of purposes. The Eurobarometers, for instance, are not academic surveys and their objectives are decidedly non-academic. The same might be said of USAID surveys. On the other hand, the World Values Surveys, though designed by individuals and organizations of varying academic commitment, are conducted primarily for scientific purposes. It should therefore not be surprising that these surveys differ substantially in their methods and the quality of the data they produce.

I begin by applauding all efforts to develop cross-national survey data bases since they have the potential to generate nomothetic rather than idiographic conclusions and insights. Consider our findings that Russia has developed a reasonably democratic political culture in the post-World War II period (e.g., Gibson 1995, 1996). Numerous referees of our various papers have asked: Compared to what? How democratic are the cultures of the most successful democracies and how does Russia compare to those? Comparison is the essence of science, so the impulse to judge whether a glass is half full or half empty on the basis of cross-national comparison is ever-present and entirely reasonable (as are efforts to compare levels of corruption across systems).

But, as with data sets like those produced by the World Health Organization, the devil is in the details.

> "Tremendous progress is being made in measurement, progress to which the rest of the field should play close attention."

Certainly, practices such as using student surveys or focus groups (or even surveys of IBM employees throughout the world) to infer country characteristics are highly dubious methodologies with well-known and obvious limitations. If one seeks to compare countries (and I recognize that other legitimate research objectives exist), representative within-country samples are essential.

Special mention should be made of focus groups, since their popularity has recently exploded. I find focus groups extremely useful, and employ them at the preparation stage of a project to get some feel for how *some* people think about issues and the language and concepts they use (e.g., Gibson 2004). But focus groups have no role in drawing inferences and conclusions about country characteristics. Social scientists must be vigilant in not generalizing the findings of focus groups beyond the handful of people participating in such discussions.

Cross-National Measurement: Validity and Reliability

The single most important detail that bedevils multi-country surveys is measurement and the most significant worry of measurement in cross-national research is language. Many of the concepts in which political scientists are interested transport relatively easily across most national boundaries (especially with the omnipresent diffusion of innovation). Language, however, is another matter. The paramount challenge for multilingual research is maintaining stimulus comparability through the translation process.

Established and effective methods exist for developing multilingual questionnaires, and I have had considerable experience with preparing such documents in as many as eight languages in South Africa (from Afrikaans to Zulu), as well as in the various languages of the European Union. The process employed is arduous, expensive, challenges mightily those with little tolerance for ambiguity, and only minimizes error rather than eliminating it.[2] But when combined with the crucial practice of using multiple indicators of each concept (see below), comparability can readily be established. More specifically, measures can be developed from which empirical estimates of both validity and reliability can be derived.

One of the least useful practices plaguing survey research (as well as many other forms of research in comparative politics) is the use of single-item indicators. For instance, few measures are of less utility than the conventional "satisfaction with democracy" question ("On the whole, are you very satisfied, rather satisfied, not very satisfied or not at all satisfied with the way democracy is developing in your country?") What does a "dissatisfied" response actually mean? Does it mean that democracy is not worth it, that authoritarianism is preferred, or does it denote that the implementation of democracy has been incomplete and that efforts to democratize must therefore be redoubled? What do respondents understand by "democracy," and what assurances exist that each respondent understands the term similarly?[3] Unfortunately, single-item indicators like this offer no means of assessing validity or reliability (even though the weak "face validity" of this item alone undermines its value).

An illustration of how valuable multiple indicators can be might be in order. In our early surveys in the Soviet Union we used an item from Rokeach's Closedmindedness Scale that asked respondents to agree or disagree with a statement to the effect that "One cannot be too careful in choosing one's friends." The item was meant to tap into the view that the world is made up of good and evil, black and white, and one must always be vigilant. However, in the context of widespread shortages in the Soviet Union in the early 1990s, the question assumed a different meaning: In order to get by, to be able to acquire the things one wants and needs (from cigarettes to shoes), one could not in fact be too careful about choosing one's friends. How did I know that this item failed as a measure of dogmatism? Through statistical analysis of its relationship to the other five indicators of dogmatism, the inadequacy of the measure became painfully evident. The use of multiple indicators of concepts is absolutely essential in all areas of the social sciences, including cross-cultural survey research.

The lesson here is simple: Established methods exist for developing multilingual questionnaires, even if it is unclear that the necessary steps of translation, back-translation, and reconciliation are implemented in all cross-national surveys. When combined with the use of multiple-item indicators, reasonable cross-national comparability can be established. Perhaps more important, rigorous indices of within-system validity and reliability can be calculated and compared.

The Special Problems of Omnibus Surveys

Omnibus surveys are particularly susceptible to problems of measurement, primarily because they seek to measure as many concepts as possible within a single survey, so as to satisfy the expectations of large numbers of collaborators or constituents. The use of multiple indicators gives priority to validity and reliability, at the expense of breadth of measures. Omnibus surveys rarely use multiple indicators for the concepts measured.

Teams conducting omnibus surveys occasionally do not include expertise in all concepts being measured, and the results can be disastrous. For instance, the World Values Survey completely mangled the measurement of political tolerance in its 1995-1997 manifestation (see Peffley and Rohrschneider 2003 for a valiant attempt to salvage the data).[4] Often the measurement of concepts is fairly complicated and advanced; the presumption that a novice can get up to speed virtually overnight is mistaken. Thus, one conclusion I reach is that cross-national surveys of narrow scope, conducted by experts on a particular topic, will likely produce more valuable outcomes than omnibus surveys.

The Problem of Transparency

Survey research seems relatively simple, but appearances are deceptive. I am often taken aback by those who believe they can design and conduct surveys without extensive training, or even by those who think they can effectively analyze survey data without training in psychometrics. It is true that large cross-national projects are typically fielded by experienced researchers, but it is difficult for me to understand how insightful analysis can be extracted from surveys without an understanding of acquiescence bias, measurement theory, or even the seemingly trivial problem of "don't know" responses.[5] I would certainly be loath to use cross-national unemployment data without carefully studying the strengths and weaknesses of such data; one might think that survey data would be treated with similar caution and seriousness.

> "Because surveys are expensive, survey research firms and organizations have strong incentives to cut expense by employing questionable survey practices."

One difficulty with assessing the quality of many large-scale cross-national survey projects is that crucial details about the survey methodology are often unavailable. Though it is difficult to publish the results from a single country survey without reporting the response rate (indeed, the *APSR* requires that response rates be calculated according to the recommended formulas promulgated by the American Association for Public Opinion Research), multinational studies are often not held to the same standards as single-country surveys (just as within-system longitudinal studies based on surveys, e.g., presidential popularity, rarely report any details whatsoever about the component surveys). If one is to use data generated by others - Transparency International, the OECD, Freedom House, the EU, or the Latin Barometer - one must be able to assess the validity and reliability of the measures and explicate and defend the overall study methodology. How many languages were used (a question even relevant to countries like Spain)? How was translation accomplished? What "don't know" options were made available to the respondents and when in the interview? Who were the interviewers and did their characteristics match those of the respondents? How was the sample selected? Was respondent substitution allowed? What was the response rate (and how was it calculated)? Response rates are particularly important, not only because of their obvious relevance to sample representativeness, but because the methods used by some commercial firms make the calculation of response rates impossible, and are therefore suggestive of more fundamental problems with the survey design. A panoply of technical details must be made available for all surveys, irrespective of whether they are single-country or multi-country projects, if we are to be able to assess the quality of the data collected.

Cost Considerations

A defining characteristic of cross-national survey projects is their cost. Cross-national survey research is probably among the most expensive methods of data collection currently employed by social scientists. Because surveys are expensive, survey research firms and organizations face strong incentives to cut expenses by employing questionable survey practices. It is little wonder that corners are cut when the cost differential between the cheapest and most expensive available surveys might involve a factor of ten or even higher. Given the large sums of money involved in survey research, it is also unsurprising that survey data are occasionally compromised. As a result, we must be vigilant in exposing the details of the methodologies of cross-national survey projects. This problem, of course, applies to single-country studies with equal force.[6]

Generalizing these Concerns

Most of my comments here pertain to surveys of representative samples of country populations. But many of the difficulties and desiderata apply as well to elite surveys, surveys of entities other than people (e.g., organizations), or even to surveys of so-called subject-matter experts. Survey research is actually more ubiquitous than many would imagine once the wide range of contexts - not just surveys of the mass public - in which samples of people are asked systematic questions is considered.

Concluding Comments

Over the last ten to fifteen years, survey research has become spectacularly reinvigorated. The old prototype of simply asking questions and recording answers has changed dramatically. Experiments embedded in representative surveys are now the rage, in part due to the strength of the causal inferences supported and in part because contextual factors can so readily be built into experiments. Research on persuasion, opinion leadership, framing, deliberation, etc., is now commonplace and has produced valuable new insights into the dynamics of opinions, attitudes, values, and behaviors. Surveys are now being fruitfully combined with other sources of data (e.g., data characterizing the environment in which the respondent is located; data on the organizations in which the respondent claims to participate). Methods of analysis have changed dramatically, answering important questions, but raising fascinating new questions as well (e.g., using surveys to assess the consequences of institutional differences, as in Hansen 1998). Tremendous progress is being made in measurement, progress to which the rest of the field should pay close attention. And the problems faced by cross-national survey researchers are no more serious or threatening than those faced by institutionalists, political economists, or any other empirical branch of comparative politics.

Thus, while it is quite legitimate to ask whether the discipline is currently learning all it can from cross-national survey research,[7] it is also appropriate to applaud the tremendous contribution survey research has made to our understanding of comparative politics. In the end, if our goal is to draw systematic conclusions and inferences about the ways in which mass politics is similar and dissimilar across various countries, no alternative to survey research exists. My hope is that this symposium (as well as other efforts currently underway) will contribute to strengthening the craft of cross-national survey research and making our findings and results more compatible across systems, and, therefore, more valuable to political science.

Notes

[1]I appreciate the comments of Robert Rohrschneider, Bill Mishler, and Art Miller on an earlier version of the article.

[2]I have occasionally found that the translation process reveals fundamental deficiencies in the English version of questions. For instance, consider the difficulty of making the following question meaningful in the context of a communist society (emphasis added):
On this card are three basic kinds of attitudes concerning the society we live in. Please choose the one which best describes your own opinion.

CODE ONE ONLY
1. The entire way our society is organized must be *radically* changed by *revolutionary* action
2. Our society must be gradually improved by reforms
3. Our present society must be *valiantly defended* against all *subversive* forces
After struggling to translate this item into Russian (and failing), it became clear to me that this question likely generates little useful information *in any language.*

[3]For this and other reasons, we rarely use some grand concepts such as "democracy" in our survey questions. Instead, our approach is to disaggregate large concepts into their compo-

nent parts (e.g., multi-party elections) and then to ask questions about the individual constituent elements.

4 Measuring political tolerance in the cross-cultural context illustrates a valuable technique for establishing comparability. Rather than asking respondents whether they tolerate groups selected by the researcher, we ask them to first identify a political group they dislike a great deal (the “least-liked group”), and then to judge whether that group ought to be extended civil liberties. Thus, the “out-group” in a society is defined by the members of that society, with the result that nominal equivalence is lost (e.g., the questions are asked about different groups in different societies), but conceptual equivalence is maintained.

5 For this and other reasons, efforts to establish and compare “point estimates” (e.g., estimates of the percentage of people who trust others) are often difficult since these estimates are influenced by a wide range of factors (e.g., question order, and whether “don’t know” is offered as an explicit option). More fruitful is the comparison of the nature of relationships among variables.

6 For instance, in France, the norm is to use quota samples in most areas of research. A true probability sample is bid by survey firms as vastly more expensive than the standard quota sample.

7 Although I should note that the questions put to us include certain assumptions and judgments about survey research that are themselves contentious. For instance, it is not at all clear to me that criticism of *The Civic Culture* impeded or undermined cross-national survey research in the period from 1963 to 1973.

Improving Data Quality in Comparative Survey Research

Debra Javeline
University of Notre Dame
javeline@latte.harvard.

The main problem with surveys used in comparative politics research is not that they are single-country studies where they should be multi-country or vice versa, but that they tend to produce poor quality data. They are often crafted by country specialists with little knowledge of survey research methodology or by survey researchers with little knowledge of the countries under investigation. The important question is: How can we conduct each type of survey research, cross-national and single-country, better?

Improved data quality can be achieved by both the occasional practitioner of survey research and the seasoned veteran. Contributions can also be made by consumers of comparative survey research. Each group has a different role to play. Below I suggest how.

The Occasional Practitioner

Comparativists who dive into survey research with no prior training or experience display too much confidence in some respects and too little in others. We are overconfident in believing that we are well prepared for a new research tool simply because we are smart and experienced using other research tools. At the same time, we lack confidence in or fail to appreciate sufficiently the importance of our work, which leads us to begin research ill prepared. One obvious solution is to commit more seriously to studying the survey research and social psychology literature and getting survey research experience prior to running surveys as principal investigators. Another less-implemented but often more practical solution is to join forces more regularly with survey research methodologists.

Look before you leap.

We as a community spend a vast amount of time learning multiple foreign languages, the history and culture of certain countries, our substantive areas of interest such as nationalism or political participation, and increasingly, statistical analysis. If our research then hinges on data collected via survey techniques, learning those techniques should be an equal if not higher priority. All these other skill sets can shine only in the presence of high-quality data.

Just as we would not recommend beginning research on electoral systems without first reading the literature or trying to run a logistic regression without first learning statistics, so we should not just sit down and write a questionnaire. Before conducting a survey, comparativists should become well versed in survey research techniques. Moreover, we should appreciate the extent of this undertaking. We should tackle the large body of literature that offers practical guidance on the basic dos and don’ts of survey research methodology, as well as the more sophisticated body of literature on social psychological theories of survey response. We should familiarize ourselves with the thousands of pre-existing survey questions that have been vetted over the decades by some of our colleagues and critiqued by others. (Survey research is the one area of academia where plagiarism is not only welcome, but encouraged. Still, comparativists should plagiarize critically, since most pre-existing questions were constructed in English for respondents in a U.S. context.) When possible, we should serve as apprentices to experienced survey researchers and see first-hand the many potential pitfalls. Acquiring survey research expertise is not as simple as picking up an introductory textbook or reading a handful of articles.

"Acquiring survey research expertise is not as simple as picking up an introductory textbook or reading a handful or articles."

If you or a comparativist you know is conducting a survey for the first time, get help.

Many comparativists are understandably reluctant to commit the time and energy necessary to acquire survey research expertise. The field of comparative politics already demands a great deal of its students, and time is at a premium. Still, since there is no shortcut to survey research training, if time-constrained comparativists want to conduct surveys without investing appropriately in the training, they should team with survey research methodologists. Only then will they produce more valid, usable, and lasting data.

The commitment to going it alone in complicated research projects needs to be abandoned. Civilization advances by specialization of function, and so should comparative politics research. Celebrate diversity in skills. Accept what you are not (a statistician, a sampling expert, a questionnaire design expert), and accept that you will never be some of these things without sacrificing other skills that are also essential in comparative politics research and that you may enjoy more and perform better. Farm out (meaning pay for) the tasks that you are not qualified to do, or take on a coauthor. If you cannot pay, you must do the latter. You should not assume these tasks are easy or less essential to your research than the skills you bring to the table.

Tapping your department's most experienced survey researcher and asking this person to review your draft survey instrument is not sufficient. In most cases, the survey researcher's time is already heavily taxed by similar favors, and although he or she may give useful feedback, time constraints virtually guarantee that the feedback will fall short of this person's standards for his or her own personal work. Working with a foreign survey research firm is not sufficient. If the firm is not your collaborator on the planned written work, then it is capable of the shortcuts typical of employees. Even if you have an ongoing relationship for a multi-wave survey, if you do not demonstrate careful attention to detail at the outset, the firm has little incentive to do more than satisfice. In short, you have to know your stuff, or you have to find a true collaborator who knows his or her stuff.

The Seasoned Survey Researcher

As the gatekeepers to the survey questions used to generate some of the most widely analyzed data in comparative politics, comparative survey researchers are in a position to advance the field by resolving some ongoing debates about measurement. Currently, a tension underlies almost every decision in questionnaire design: Do we keep the time-tested question in order to document trends and conduct time-series analysis, or do we admit that the time-tested question does not measure what it claims to measure, confuses respondents, translates poorly to other languages and cultures, or has some other flaw that merits abandoning it and asking a more valid question? Very often, path dependence wins over innovation. Unfortunately, however, the victory may saddle us with lousy survey questions in perpetuity and thus not serve to advance our field.

Validate, don't proliferate.

Certainly, many of us do reject older survey questions in favor of newer ones, but this practice results in its own problem for the field: Since we make these decisions independently and often without knowledge or consideration of the decisions of our colleagues, comparative survey research (and political science and social science more generally) have very few agreed-upon measures. In order for us to become a more mature discipline, we must address similar research questions with the same measures and build on previous research so that knowledge is cumulative. We need to agree on how to measure widely discussed concepts such as trust, efficacy, individualism, nationalism, liberalism, etc. before we can effectively move on to the bigger tasks of theory-building and testing.

Some of these issues are being addressed by political scientists engaged in "basic" survey research to introduce new measurement techniques, such as vignettes that build on time-tested survey questions (King et al. 2004). However, from a validity standpoint, the vignettes are only as useful as those original survey questions. If they are built around survey questions that are poor at the outset, we still cannot be sure of exactly what we are measuring. We need to address whether the efficacy question really measures efficacy, the trust question really measures trust, and so on.

"You are not accountable for data you did not collect, but you are accountable for using the data uncritically."

For example, take the following widely-used survey question, "Generally speaking, would you say that most people can be trusted, or that you can't be too careful dealing with people?" Critics challenge whether this question truly measures a personality trait called "interpersonal trust", as its proponents claim, or whether the question measures the objective reality of living among trustworthy or untrustworthy people. Or take the question, "Do you think that quite a few of the people running the government are crooked, not very many are, or do you think hardly any of them are crooked?" Critics challenge whether, in a cross-national context, this question truly measures trust in government or perceptions of government corruption, or whether the word "crooked" translates so poorly from the English as to make cross-national comparisons meaningless. Or take the standard political efficacy question, "Public officials don't care much what people like me think." Critics challenge whether this question measures the perceived responsiveness of the political system, as intended, or whether it also captures a personality trait called "self-efficacy" and so confounds interpretation. Moreover, they ask: Who are "people like me"?, and does the question work differently for respondents with different reference groups based on race, profession, gender, class or some other characteristic? Resolving these debates and providing evidence of which survey questions are the most valid measures of commonly discussed concepts is one of the most valuable contributions that experienced comparative survey researchers can offer the comparative politics community.

Take seriously the findings from methodological experiments.

The debates involve not only questions of content but also questions of form. Most experienced comparative survey researchers know the methodological literature and do not dispute various findings about response bias and other survey error. Nevertheless, they often fail to use the findings as a guideline for questionnaire construction. For example, we know that "attitude statements" (those that require respondents to answer "agree or disagree," "yes or no," or "approve or disapprove") are notoriously prone to acquiescence bias, yet the ease of analysis and path dependence on previously used questions encourages many survey researchers to ignore the dilemma entirely, blanket their questionnaires with attitude statements, and analyze the resulting data uncritically (Javeline 1999). Creators of large, publicly used data sets should either implement the recommendations that follow from methodological experiments or at the very least include information on potential response biases and other survey error in the data documentation.

Moreover, experienced comparative survey researchers should interpret methodological findings to the rest of the comparative politics community and make appropriate recommendations for future research. For example, we know that there are "tone of wording" effects in survey questions and that responses differ to substantively similar questions like whether to "forbid" or "not allow" public speeches by a disliked group (Schuman and Presser 1996). For methodologists, this noteworthy finding represents the endpoint of their research. For comparativists interested in asking such a question, this finding is only the beginning. The remaining, more pressing dilemma is: Which phrase – "forbid" or "not allow" - should be used in the questionnaire? Which circumstances merit using one tone over the other? Guidance from experienced survey researchers should encourage greater awareness of response biases and greater consistency in measurement across surveys.

Be truthful in advertising

Comparative survey data often contain weaknesses known only to the collector of the data. Perhaps flaws in the sampling design limit the generalizability of findings, or perhaps controversial translation decisions make cross-national comparisons inappropriate. Concealing such weaknesses encourages overconfidence in findings and ultimately magnifies skepticism in the larger political science community of all comparative survey data.

Comparative survey researchers should instead make available honest critiques of their surveys' weaknesses, copies of all questionnaire translations, and justifications of sampling, translation, and other decisions. This documentation should be as easily accessible as the dataset and codebook. Even surveys with weaknesses may still represent tremendous contributions to our field. Many of us are working in post-authoritarian and/or developing countries with new and limited survey research infrastructure, and we are some of the first to try to collect survey data systematically. Full disclosure of pitfalls and tough decisions is not an admission of failure; it is a sign of early success in a maturing field. It will help consumers of the data interpret findings accurately, and it will help our comparative survey research successors collect better data in the future.

The Survey Research Consumer

You need not be a practitioner of survey research to help improve data quality. Just as consumers in other markets can alter goods by demanding more from producers, so too can consumers of cross-national and single-country survey data. Educate yourself about survey research methods enough to have a healthy cynicism about the data you use. Be honest about the data's strengths and weaknesses. Expose flaws when you see them. You are not accountable for data that you did not collect, but you are accountable for using the data uncritically. Moreover, your vigilance will lead the survey research practitioners to be even more careful in their future pursuits, because they will know that their audiences are capable of greater scrutiny.

If you feel ill-equipped to judge the merits of survey questions, sampling design, and other aspects of survey research that produced the data on which your research depends, the above advice for the occasional survey research practitioner is relevant: Take on a coauthor, or at the very least, pay for some methodological assistance. We have come to accept as normal practice paying for methodological assistance or teaming up with a methodologist when it comes to statistical analysis. We need to broaden this view and encourage the same behavior when it comes to working with survey data. If we draw on our respective strengths and collaborate more regularly, rather than trying to be jacks-of-all-trades, we will spend our time more efficiently and produce more and better research.

Conclusion

Survey research is often considered "soft" or even "easy" by many comparative political scientists. Courses on social science methods focus on the supposedly more difficult stage of data analysis, and the process of data *collection* is often ignored or mentioned only briefly. Instructors implicitly assume that comparativists will analyze only secondary data or that, if they need original data, they can quite easily acquire the survey skills on their own.

In reality, many of us are going beyond secondary data analysis and collecting original survey data without sufficient training, and we are finding that data collection is by no means easy. Comparative survey research is not a science, and the lack of clear right answers makes it difficult rather than easy. Without strict formulas or even a general consensus on the most advisable procedures, the room for error is huge. Moreover, once we collect the data, there is no going back. If a question turns out to be flawed, the principal investigator and the rest of the comparative community are stuck with it. High-end statistical manipulations may help, but they also may not.

Acquiring high-quality data at the outset is the best assurance that we can appropriately test our hypotheses and better understand comparative politics. Those of us who regularly collect such data should be more proactive in setting standards for the field. Those of us who need help in collecting or using such data should seek it. Survey research requires the anticipation of a wide range of perceptions and biases and is thus an inherently collaborative type of research. Collaborative data collection and co-authorship should be embraced by the field.

Note:

Complete citations for this issue are online at http://www.nd.edu/~apsacp/backissues.html.

What Has Comparative Politics Accomplished?

Gerardo L. Munck
University of Southern California
munck@usc.edu

Richard Snyder
Brown University
Richard_Snyder@brown.edu

The following feature is excerpted from the forthcoming volume, *Passion, Craft, and Method in Comparative Politics*, by Gerardo L. Munck and Richard Snyder. The authors conducted in-depth interviews with fifteen leading scholars in the field of comparative politics: Gabriel A. Almond, Robert H. Bates, David Collier, Robert A. Dahl, Samuel P. Huntington, David D. Laitin, Arend Lijphart, Juan J. Linz, Barrington Moore, Jr., Guillermo O'Donnell, Adam Przeworski, Philippe C. Schmitter, James C. Scott, Theda Skocpol, and Alfred Stepan. These scholars reflect on their intellectual formation, their major works and ideas, the nuts and bolts of the research process, their relationships with colleagues, collaborators and students, and the evolution of the field. The material excerpted here addresses the achievements and shortcomings of comparative politics and is taken from four of the interviews.

ROBERT A. DAHL

Q: What are some of the main achievements of comparative politics over the last 50 years?

A: Over the last 50 years, there has been an enormous gain in the quality and quantity of knowledge that we have in the field of comparative politics. This is a very positive and impressive change. I don't know that it's a bigger change than has occurred in other branches of political science, but it might well be, especially since comparative politics started with a rather narrow base. When I was a graduate student at Yale in the 1930s, Herman Finer's *The Theory and Practice of Modern Government* (1932) was the main text in comparative government courses. That text concentrated on the United States, Britain, France, with perhaps a little bit on fascism and the Soviet Union. At that time, comparative government was a very narrow field with a primarily historical, descriptive, and institutional content. There was a lot of valuable information there, but it was not theoretical, at least not in the sense of theory as we mean it today.

Overall, the field was very Eurocentric. There was no study of Japan or China, for example. There may have been some people who wanted to study the Soviet Union. But until the outbreak of World War II, our horizons were really quite limited. The big thing going on in the world was the New Deal. There were very few people in the field who had mastered the languages, even Russian, that were required in order to understand non-European countries. Even Latin America - a rich treasure house of experience nearby - was not within our ambit. I don't think we studied Canada. It was all very parochial. The expansion outward of the study of comparative politics beyond a handful of European cases, the growth of often quite good world-wide data, and the development of methodologies and analytical techniques have fundamentally altered the field in a positive way. We know a lot more about political parties, constitutions, and about broader things like regime breakdowns and transitions. Fifty years ago we knew almost nothing about these things. We know so much more about electoral systems, their consequences, and how they function than we did only a few years ago. The growth of knowledge about these important issues is heartening and positive. I don't want to exaggerate and say we are anywhere near achieving a final, conclusive body of knowledge; we never will be, but we know so much more. I think the study of comparative politics and comparative democracies, specifically, might be the most promising part of political science today.

Q: What are some of your major dissapointments with the field of comparative politics?

A: I think it's appalling that at this late date we are still struggling with how to conceptualize and measure democracy. I find the continuing debates about what we mean by democracy - I mean this and he means that - depressing.

Q: Why do you think it has taken so long to come up with a satisfactory definition of democracy?

A: Part of the reason is that a satisfactory definition has to respect the history of the term. You don't want a definition of democracy that makes the Greek city-state undemocratic. You have to accommodate that. At the same time that a satisfactory definition of democracy has to respect the history of the term, it also has to be able to accommodate the evolution in its meaning. And a satisfactory definition also has to be formulated in a way that allows you to measure it. This requires judgments and rankings based on judgments. For example, you need to judge how free speech is in Peru. This is not like reading a scale on a thermometer, but with a large degree of observer agreement on such judgments, you can put some trust in them. These are all demanding requirements that are rarely met.

> "... perhaps we've not only failed to progress in the study of power, we've actually gone into reverse."

Q: Do you have any other disappointments with the field?

A: I'm enormously disappointed that the study of power and the conceptualization of power have made no progress that I can detect since Harold Lasswell and Abraham Kaplan's book in 1950, Jim March's work (1955,1956,1957) and my early efforts (Dahl, 1968) Power is such a central concept, and Jim March and I had hoped a vocabulary would evolve that would allow for observation, comparison, and the accumulation of information. We also hoped a precise and discriminating language for studying power would evolve, along the lines of what Lasswell had tried to develop and what I tried to develop for political analysis (Dahl, 1963). These expectations turned out to be highly optimistic. Now, fifty years later, I see people use the word and concept "power" as if we were back where we started. Even elementary distinctions going back to Max Weber - such as the distinction between power and authority, or legitimate power - seem to have been forgotten. So perhaps we've not only failed to progress in the study of power, we've actually gone into reverse.

Also, very few people study power today. I don't know what explains this. Perhaps the requirements for studying power in a way that we would now regard as methodologically sound and reasonable outstrip our capacities for definition and measurement. Maybe the problem is that we don't have good ways of measuring power, so the people who would be likely to study it know that the methodological requirements are just too daunting.

JUAN J. LINZ

Q: Several scholars (Lindblom 1997, Geddes 2003) have recently questioned whether comparative politics has generated cumulative knowledge? What do you think about this issue?

A: I think there are several different areas of research in which we have quite a lot of learning and cumulative work. One example of learning can be seen in the research on consociational democracies by Arend Lijphart (1968,1977) and others That work persuasively challenged the old prevailing notion that a majoritarian, two-party model of democratic politics worked far better than a multi-party system. Now we know that a number of democracies with multi-party systems have actually worked very well. That's an insight we didn't have when some of the literature on political parties started. Similarly, the research on corporatism, starting with Phillippe Schmitter (1979) and Gerhard Lehmbruch (1982), taught us that a close integration of interest group politics and political parties was not necessarily a bad thing for democracy. I think the comparative literature on transitions to democracy has taught us something about how to make a transition possible. In particular, we have learned that transitions take place within the institutional framework of the previous regime, can be negotiated, and are not necessarily violent breaks with the past. To take another example, work by people like Giovanni Sartori (1994), Rein Taagepera and Matthew Shugart (Taagepera and Shugart 1989, Shugart and Carey 1992) on the consequences of different electoral laws has also generated lots of solid knowledge. Finally, comparative work on elections using survey research has produced a lot of cumulative knowledge about the relationship between social variables, like class and religion, and voting behavior. The problem is that our knowledge about something like voting behavior is cumulative only until you have a political earthquake. Then some of the relationships among the variables no longer hold. For example, the library I have in my basement on Italian politics and voting was very cumulative from 1948 to the 1990s. But then the Christian Democratic Party disintegrated, Berlusconi and the Northern League emerged, and the whole party system changed. This means that to understand Italian elections you have to start from scratch, well, not from scratch, but you have to start anew. Previous research cannot be the only basis for analysis. By creating these discontinuities, Italian politicians were being cruel in a sense to the social scientists that had spent decades working on Italian parties and elections.

More generally, all the knowledge generated by research on voting behavior in advanced industrial democracies is getting dated. This is partly because the working class, which used to represent some 30 to 40 percent of the population, has shrunk dramatically in many places. The old linkages among variables like working class identification, trade union membership, and participation in labor, social democratic, or communist parties have been weakened. And you increasingly find a much more homogenized, "middle class society" in these countries. As a result, the old loyalties to parties have eroded. So, a worker who, in the past, would have said "I am a worker, therefore I am a union member, and therefore I have to vote social democratic or communist" today might say, "I am a worker, but I have a summer home on the Mediterranean coast which I rent to tourists, and the social democrats propose to raise taxes on my second house." So, that person, even though he is still a worker, may vote against the party he would have voted for in the past. And voters today are much freer in some ways. For a long time, the Italian voter saw the Christian Democrats as his protection against the threat of the communists coming to power. So, no matter what misgivings he may have had about the Christian Democrats and the corruption that characterized their governments, he voted Christian Democrat.

Today, no parties are seen as a serious threat, and Italian voters thus feel much freer to vote for whichever party they think best fits their interests. This makes it infinitely more difficult to predict voting behavior in Italy. To take another example, decades ago, whenever I met a Dutchman, I only had to ask him two or three questions - Are you a Catholic, a Calvinist, or a non-believer? What's your occupation? - and I knew how he would be voting, because 90 percent of the voters with certain social characteristics voted for a particular party. That's not the case any more, which obviously makes the study of political parties more difficult. All these types of changes set limits on our ability to generate cumulative knowledge about politics.

> "... the problem is that our knowledge about something like voting behavior is cumulative only until you have a political earthquake."

Q: What are some of the areas of research in comparative politics where you think our knowledge is especially limited and in need of improvement?

A: I think we know much too little about political leadership and the quality of political elites. We know that political elites usually have higher education, come from a certain background, know foreign languages, have studied abroad, and so on. But we do not know why some leaders are more creative and more committed than others, and why some leaders are real crooks. As Schumpeter (1942) notes at the very beginning of his theory of democracy, in order to have a working democracy you need a pool of qualified people who are committed to public service. Why do some societies have these people but others don't? And why do some countries produce creative business elites? In today's *New York Times* there is a story about the Hyundai empire. Why did the Hyundai empire emerge in Korea and not in Argentina or somewhere else? These are things we don't know. The number of things we know so little about is startling. There is so much work to be done.

SAMUEL P. HUNTINGTON

Q: What are the most important achievements of the field of comparative politics?

A: If you go back to the 1950s, scholars in comparative politics performed a very useful service in rapidly studying the politics of developing countries. Through World War II comparative politics really was nothing but the study of the major European powers and the United States. If you look at any textbook of comparative politics from back then, you have five chapters: one on the United States, one on Germany, one on France, one on Britain, and so forth - that's it. It seems to me that with all its shortcomings, the literature on political development that emerged in the 50s, 60s, and 70s was very broadening and constructive. Also, comparative politics has become more sophisticated in its methods of analysis; and I'm all in favor of sophistication and methodology when it is useful, and in many cases it is useful. We now have increasing quantitative data on political development and on political variables and non-political variables that are relevant to politics. For example, Karl Deutsch back in the 1960s put together the *Handbook of Political and Social Indicators* (Russett et al., 1965) - which was a very useful compilation of data. Now more data are available, and Ronald Inglehart, with the World Values Survey, which has limitations stemming from the fact that he has to rely on some not terribly sophisticated polling outfits in Third World countries, has compiled a very useful source of quantitative information on the values of people around the world (Inglehart 2003).

> "There have certainly been a wide range of propositions - empirical generalizations - which have stood the test of time."

Q: I'm struck by the fact that you haven't mentioned any big theoretical contributions or breakthroughs as major achievements.

A: Theories come and go. The subfield has gone through phases - back in the 50s structural-functionalism was the big thing - Gabriel Almond and James Coleman's *The Politics of the Developing Areas* (1960) for example. Theoretical frameworks like that come and go - they may be useful for a period of time. In responding to your question about the major accomplishments of the subfield, I was trying to think of things that made a more lasting improvement. There have certainly been a wide range of propositions - empirical generalizations - which have stood the test of time. One example is the positive relationship between economic development and democracy. Marty Lipset set it forth in a rather crude form back in the 1950s. It's been refined, and dozens of scholars have built on his initial article. But overall his proposition holds up. Propositions about the impact of social and economic change on social and economic equality have also been enduring. In my own work (Huntington, 1968), I showed - and other people have supported the same conclusion - that political instability is not a product of poverty, it's a product of people getting *out* of poverty. I think the work of Donald Horowitz (1985, 2001) on ethnic conflict is also a very considerable contribution. At the conjuncture between comparative politics and international relations the so-called "democratic peace" proposition was a major contribution which, despite all the debate about it, holds up pretty well.

Q: What are the major shortcomings and failings of the subfield over the last 40 years or so?

A: Political science - including the study of comparative politics - is heavily influenced by other disciplines. The principal external disciplinary influences change from time to time. During the past several years, economics has been dominant. Before that it was sociology, and certainly Almond and others were heavily influenced by Talcott Parsons. But even before that there was the whole psychological approach to politics - for example, somebody like Harold Lasswell in the 30s and 40s tried to look at politics in terms of its psychological dimensions. It seems to me it would be useful to go back and revisit that earlier work on psychological approaches to politics, in part because there's always the tendency - in the laudable desire to produce generalizations - for people to look at issues where you have a large number of variables that you can study, compare, quantify, and so on. This tends to omit the decisive role which political leaders play. The serious study of political leadership has been a great deficiency in recent political science.

In addition, 30 or 40 years ago, there used to be a very considerable literature and a lot of work done on political socialization. It might be time to go back and study political socialization, which basically means how people's political values develop. A lot of the literature 30 or 40 years ago was looking at children, asking from where they get their ideas about politics, about political leaders, the presidency, and so on. We're in a very different era now, and it would be useful to go back and look at political socialization in that context but also in a broader context about how political values change. If culture is important and there are some cultural systems - systems of beliefs and attitudes - that are conducive to economic and political development and other cultural systems that are not, the next question is how do you go about changing a culture? How do you change people's political and social attitudes, beliefs, and assumptions? That's a very important question on which I don't think much relevant work has

> "The serious study of political leadership has been a great deficiency in recent political science."

been done. One can think of examples of things that do change values. A traumatic event will change people's values. For example, Germany and Japan in the 1930s were the two most militaristic countries in the world, but the trauma they suffered in World War II turned them into two of the most pacifist countries in the world. That was a real change in values. Economic development changes people's values. I think that's very clear and can be seen in Inglehart's work (1990, 1997) on the development of materialist values and the shift to post-materialist values. But if you want to change values in order to *produce* economic development, that work is not much help.

ADAM PRZEWORSKI

Q: If you look at where the field of comparative politics was 30 years ago and where we are now, what are the main things we have learned?

A: Let me preface my answer with one caveat. I think some of the best research in comparative politics is done these days by economists, so I will include them in my answer. Daron Acemoglu and James Robinson, Alberto Alesina, Torsten Persson and Guido Tabellini, and many others do excellent work in comparative politics. They don't know enough about politics, particularly about institutions, but they address central questions and get provocative answers. With that inclusion, yes, I think there has been a tremendous accumulation of knowledge.

What have we learned? Ever since Maurice Duverger's (1964) and Douglas Rae's (1969) seminal books, we have learned a lot about the consequences of electoral systems. Gary Cox's book, *Making Votes Count*, (1997) is the latest example of it. We know how the electoral systems interact with social cleavages to produce parties, how they affect the distributions of votes, and so on. We've learned a lot about coalition formation and cabinet formation; there is a formal and an empirical literature on these topics. We understand much more about the legislative process. We've learned a great deal very rapidly in the last few years about ethnic conflict and ethnic peace. We have learned that most of the time ethnic groups live together in peace, and perhaps we are beginning to understand some mechanisms that explain this finding. Finally, I think we understand much more about the processes of regime transitions. I could go on.

More broadly, one test of the advances we have made is that when a student raises a topic with me, most of the time I can say "okay, read this, read that, here is the literature that says this and that." On various topics, the conclusions do not converge. But at least there are bodies of literature on a variety of topics.

Q: Are there any topics on which we have not made significant advances?

A: We still do not know why and when people with guns obey people without them: the determinants of civilian control over the military. We still don't understand political parties very well. This is truly an important topic, which we have neglected. We don't understand why parties come into existence, what mechanisms hold them together, and what the glue of party discipline is. Though we have learned a lot in general about authoritarianism, I also think we know disastrously little about the structure of dictatorships. We tend to think that formal institutions are just a window dressing and yet my student, Jennifer Gandhi, found that they matter for the survival of dictators, for policies they pursue, and for the outcomes they generate. I think we are not doing well with globalization. I've written something on it recently, so I was forced to read the literature. I found it deeply unsatisfactory. In particular, the political consequences of globalization are poorly understood. I think the problem, in part, is that we need some kind of methodological breakthrough in this area of research. The methods that are currently used just don't do well enough. The findings are disparate, and most of them are based on statistical methods that assume that observations of particular countries are independent. So it is hard to believe the statistical findings. This is a big, important topic. Somehow we are going to have to start thinking differently and pay more attention to the sort of methods that would be appropriate for studying this issue. Generally, to a large extent because of the availability of data, we know more about the OECD countries than about the less developed ones. But this gap is rapidly closing.

"I think some of the best research in comparative politics is done these days by economists."

Q: You have emphasized the methodological difficulties of addressing complex questions in a

rigorous fashion. Another reason why progress on such questions might not be made is that comparativists simply fail to pose big, interesting questions about politics in the first place.

A: What is it that we are not asking? Certainly, we are not asking “So what does all that we do know add up to?” But we also fail to ask several questions that are researchable with the methods we have. What determines the access of moneyed interests to politics? What is it about our democratic institutions that makes people feel politically ineffective? Why is it that these institutions perpetuate misery and inequality?

There is a saying in my native language, “It is not the time to cry over roses when forests are burning.” And as I talk to people in Argentina, France, Poland, or the United States, I hear that they are burning. People around the world are deeply dissatisfied with the functioning of democratic institutions, in the more as well as in the less developed countries. They see politicians as serving the interests of the rich, of multinational corporations. They cannot understand why democratic institutions seem to be impotent in reducing glaring and persistent inequalities. They feel that political parties do not serve as a mechanism of transmission of their values and interests. They perceive that important decisions are made by institutions, often international, over which no one has control. The danger is that unless we keep asking such questions, we leave the answers to demagogues of different ideological stripes.

The entire structure of incentives of academia in the United States works against taking big intellectual and political risks. Graduate students and assistant professors learn to package their intellectual ambitions into articles publishable by a few journals and to shy away from anything that might look like a political stance. This professionalism does advance knowledge of narrowly formulated questions, but we do not have forums for spreading our knowledge outside academia; indeed, we do not communicate about politics even among ourselves. It has been decades since professional journals - “professional” is what they are called - published essays on “What is wrong today with the United States, with democracy, or what not?” or on “How to make the world better?” We have the tools and we know some things, but we do not speak about politics to people outside academia.

“People around the world are deeply dissatisfied with the functioning of democratic institutions, in the more as well as in the less developed countries.”

Note:

Complete citations for this issue are online at http://www.nd.edu/~apsacp/backissues.html.

A Review of the World Bank Data Base on Political Institutions

Edurne Zoco
University of Notre Dame
ezoco@nd.edu

*Developed by Thorsten Beck, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh.

The Database of Political Institutions (DPI for short) compiled by the Development Research Group of the World Bank represents a major effort in collecting cross-regional quantitative data in the field of comparative political economy and comparative political institutions. The strengths of the database are its extensive country coverage, the creation of new variables, and a detailed variable construction On the other hand, the DPI contains some important weaknesses that future users should not overlook, namely inaccurate data collection, some coding and measurement errors, and the definition of democracy that was used. These flaws call for the need to continue improving existing databases to turn them into valuable foundations for empirically testing research in the field.

Since its first publication in 1997, the DPI has undergone several revisions and changes in variable definitions. The last update, named DPI 2000, extends DPI's coverage through 2000. The database covers 177 countries between 1975 and 2000 and contains 109 variables. The unit of analysis is a given country during a particular year. The sample includes data from 49 Asian and Middle Eastern countries, 43 Western and Eastern European countries, 48 African countries, 30 North American and South American countries, and 7 Oceanic countries. The core of the dataset is composed of variables related to the political system and to electoral rules. There are 19 chief executive indicators (e.g. number of years in office, existence of formal constraints on the term of office), 51 party variables in the legislature including dummies (e.g. government fractionalization, opposition fractionalization), 9 legislative variables (e.g. average age of parties, majority percentage), 14 electoral variables (e.g. legislative and executive index of political competitiveness, House and Senate mean district magnitude), 9 checks and balances indicators (e.g. number of veto players, polarization), and 5 federalism variables (e.g. existence of autonomous regions, existence of directly elected sub-national governments).

Unlike other datasets such as the Freedom House or Polity (III and IV), DPI's variables are based on "objective" institutional characteristics and, for the most part, are disaggregated, with the exception of two indexes (the index of legislative and executive political competitiveness). Thus, several variables in the database are dummies. This is clearly an advantage for researchers looking for disaggregated data. It does not represent a disadvantage for scholars looking for aggregate variables given that it does not impede a later aggregation of the data customized to the interests of researchers.

The recent publication of databases such as the DPI is a response to the increasing scholarly interest in the study of and advancements in the fields of political economy, political institutions, and democracy. To date, research in political economy, political institutions, and the determinants of institutional change has been - and is - contingent on the availability of empirical data to test hypotheses. Data availability is a major issue in comparative politics, especially in cross-regional analysis and, in this sense, the DPI advances the field through the empirical testing of new research questions and hypotheses.

In addition to its extensive coverage, DPI's major contribution is the creation of new variables and the improvement and extension of several already existing measures (e.g. checks and balances, government turnover). Equally important, the process of variable construction is very detailed in most cases. It is reassuring for data quality purposes that authors have explicitly stated the rules employed during the process of variable creation and coding as well as the dilemmas and shortcomings presented by some variables in the database. The authors themselves recognize some difficulties with the construction of some variables and they have attempted to solve some critical points in the last database update.

Nevertheless, the DPI contains some conspicuous flaws that need to be addressed. One of the database's major weaknesses results from inaccurate data collection. Some miscodings exist not only among some authoritarian countries and new democracies (where traditionally data collection is more intricate) but also among several European countries. More troubling is the fact that several inaccuracies in data collection come from variables where information is easily available (e.g. web pages of national congresses). For example, in Spain the Basque Nationalist Party (PNV) is not coded as a nationalist party; the sum of seats of government and opposition parties does not equal the total number of seats in the parliament; and in the 1977 election, the Socialist Party (PSE) is included in the database as both the first and the third opposition party.

Furthermore, the measurement and coding of "party platforms" of the chief executive party and "polarization" is confusing and unconventional. Previous measures of "party platforms" (such as the Party Manifestos Project) as well as other measures of party ideology have traditionally used a one-to-ten scale (e.g. Castles and Mair 1984, Huber and Inglehart 1995). However, in the DPI, the variable measuring "party platforms" (EXECLRC) is coded with letters and numbers: (L) for the Left, (R) Right, (C) Center, and zero for those cases that do not fit into the left-right scheme (30 percent). Such categorization presents additional problems for estimating "polarization," since that variable is calculated using EXECLRC. It is unclear how the variable "polarization" (a numerical variable that goes from 0 to 2) was calculated from a categorical variable. The codebook for DPI 2000 does not offer further information about the procedure for converting the letters into numbers. Moreover, "polarization" has been traditionally calculated as a standard deviation or by subtracting the party-system mean on the left-right continuum from each party's left-right score, and multiplying it by its share of seats in the last national election celebrated in every country, or using variations of the same formula (Coppedge, 1998). In this database, however, "polarization" is defined as the "maximum polarization between the executive party and the four principal parties of the legislature" (Keefer, 2000: 20). This generates large discrepancies between polarization estimates calculated using the traditional formula and the DPI values. Finally, another puzzle is why those political systems where a party has an absolute majority of the seats have a polarization equal to zero (e.g. Spain 1982).

Regarding the creation of new variables (such as the index of electoral competitiveness), the main concern is whether the actual coding is the most appropriate to capture nuance and variability across regions. DPI's index is built upon Ferree and Singh's (1999) index. The index was originally built to classify only African countries on a six-point scale. The application of that index outside the African region raises some significant validity issues. Although, in widening the coverage of the index, the DPI includes a seventh category, it might be argued that this scale is rather inadequate for a large cross-regional sample with higher levels of competitiveness. The highest level of competitiveness in the electoral competitiveness scale, level 7, includes those countries where "the largest party received less than 75 percent of the seats." Following such classification, 46.3 percent of countries in the database are coded as "7" and 57.2 percent are coded as "6" or "7" making the variable distribution strongly skewed to the right. The distribution is even more skewed if the African countries are excluded: 57 percent are coded as "7" and 65.5 percent as "6" or "7". Moreover, the actual coding overlooks remarkable switches on electoral competitiveness in several individual countries. Thus, to name a few, it classifies Brazil as "7" before and after the democratic transition of the mid-80s, Peru as "7" before and after the self-coup of 1992, and Mexico as "7" until 1993 and "6" after this date - when most Latin American scholars usually emphasize the increasing openness of the Mexican political system in the 1990s. This index also presents low correlation with traditional measures of competitiveness such as the "effective number of parties" (ENP). The ENP has been extensively used to calculate the extent of party competitiveness in a system. The DPI database treats as "7" countries with radically diverse measurements of ENP. The authors coded as "7" (the greatest level of competitiveness) countries with radically different effective number of parties, such as Venezuela 1983 (1.92), Uruguay 1994 (3.17), Ecuador 1986 (7.55), and Brazil 1994 (8.13), to continue with Latin American cases. DPI 2000 does not seem to have corrected this issue. It has added two scale points, although just for intermediary scale positions (5.5 and 6.5). Moreover, the construction of the index also undermines the validity of other new variables (i.e. CHEKS) that were created using the index of electoral competitiveness.

Another problem is the definition of democracy used in the database. A country is considered democratic when the DPI's index of electoral competitiveness is greater than 6. This is an unusual criterion for democracy. Any country is classified as democratic as long as multiple parties won seats, even if the largest party received more than 75 percent of the seats. Furthermore, when a country has an index lower than 6, the country is considered autocratic and the autocratic system is as old as the executive's years in office

(YRSOFFC). This produces some strange results in countries where autocratic regimes had different executives (e.g. Russia, Poland). For instance, in 1975 Poland had been under an autocratic system for five years and in 1982 it had been autocratic for two years. This is a rather unusual measure of the longevity of democratic and authoritarian regimes and, more importantly, it conveys both the longevity of democratic regimes and the longevity of autocratic executives under the same variable. A democracy is as old as the oldest democratic government while an autocracy is as old as the number of years in office of the latest executive.

Finally, the large percentage of missing observations in some variables cannot be overlooked. Some conspicuous cases are the five federalism variables and "candidate selection" for which only 2 percent of the cases are valid or applicable. Conclusions derived from the analysis of these variables should be therefore accepted with caution.

In sum, the DPI database is an important step toward the creation of the necessary databases for our field. It represents a major effort in collecting cross-regional quantitative data in the field of comparative political economy and comparative political institutions. However, it has not yet achieved the level of data accuracy and confidence that sound quantitative research requires. Notwithstanding the fact that few databases match the great geographical and temporal coverage achieved by the DPI and fewer still include such a large array of political and institutional variables, extensive revisions must be carried out for this dataset to excel as a tool for social researchers interested in empirical cross-regional research.

Data:
http://www.worldbank.org/research/bios/keefer/DPI2000_distributed.zip
Codebook:
http://www.worldbank.org/research/growth/pdfiles/DPIvariables.pdf

Other Datasets

Dataset on Price Liberalization, Privatization and Institutional Reform in 18 Postcommunist Countries (1976-2000)

Petia Kostadinova
(University of Florida, Gainesville)

This dataset contains monthly and yearly measures of three aspects of postcommunist economic reform, price liberalization, privatization, and institutional reform. The respective indexes capture decisions by policy-makers with respect to the three reforms. The measures are based on the stage model of policy-making, which distinguishes among agenda setting, policy formulation and decision-making. Weights were assigned to government actions according to their intended effect on reform levels. Events of higher stages in the policy process were assigned bigger weights. The countries included in the dataset are Albania, Belarus, Bulgaria, the Czech Republic, Estonia, Hungary, Kazakhstan, the Kyrgyz Republic, Latvia, Lithuania, Moldova, Mongolia, Poland, Romania, Russia, Slovakia, Turkmenistan, Ukraine and Uzbekistan.

Data and codebook:
http://plaza.ufl.edu/petiak/datasets.htm

CP Dataset Award

The Award Committee agreed unanimously that the award should go jointly to Michael Bratton (Michigan State University), E. Gyimah-Boadi (Center for Democratic Development, Ghana) and Robert Mattes (The Institute for Democracy in South Africa), for their pioneering work on the Afrobarometer. We think that this innovative new data set project meets the goals of the award in every way. We were particularly impressed at the development of networks of national partner survey researchers in many countries that have never had previous social scientific surveys of political and social attitudes.

The substantive contents of the surveys are of great interest, not only to scholars, but to NGOs, policy advocates, donor agencies, and all concerned with African democracy.

Although the first round of surveys in twelve countries was only completed in 2001, many scholarly papers, as well as a book by the Principal Investigators, have already been geberated. A second round of surveys is just finished. Moreover, this project is a model of transparency in releasing full information on the work in progress, detailing the procedures, including transparency about funding and associated researchers, as well as with timely deposit of data through archives as well as their own website. For all these reasons we think that the Data Set Award's recognition of the scholars responsible for the Afrobarometer is well deserved.

Committee: G. Bingham Powell, chair (University of Rochester), Pippa Norris (Harvard University), and Chris Way (Cornell University).

Two Luebbert Awards to Isabela Mares

Best Book in Comparative Politics: *The Politics of Social Risk: Business and Welfare State Development*, New York, Cambridge University Press, 2003

Isabela Mares's book uses original data to show how different structural attributes of firms create different preferences for social insurance from the late 19th century forward. Some firms support social insurance, others oppose it. This creates the potential for political alliances across class, factor and sector. In showing conflict among firms, Mares pushes the "comparative capitalism" literature towards understanding the political processes and distributional consequences of the compromises that were reached in earlier periods and the way these shape production strategies which in turn continue to influence policy processes today. We applaud her originality, her methodology, and her ability to build on existing discussion in thinking about the underpinning of business preferences regarding social policies, to both draw upon and extend our understanding.

Committee: Peter Gourevitch, chair (UCSD), Victoria Murillo (Columbia University), and Frances Rosenbluth (Yale University).

Best Article in Comparative Politics (2003):

Isabela Mares, "The sources of business interest in social insurance: sectoral versus national differences", *World Politics*, 55: 229-258.

Committee: Peter Swenson, chair (Yale University), Stephen Hanson (University of Washington), and Beatriz Magaloni (Stanford University).

Sage Paper Award

Best APSA paper in Comparative Politics 2003:

"The Political Salience of Cultural Difference: Why Chewas and Tumbukas are Allies in Zambia and Adversaries in Malawi" by Daniel Posner (Department of Political Science, UCLA).

This paper proposes that the political-relevance of these cultural divisions depends on the relative sizes of the groups themselves. The paper, elegantly but simply written, so solves a fascinating research puzzle, using original field research with well-thought-out case selection.

Committee: Mike Lewis-Beck, chair (University of Iowa), Scott Mainwaring (University of Notre Dame), and Eva Bellin (Hunter College).

CP Section Business Meeting

The annual business meeting of the Comparative Politics Section will be held on Friday, September 3, from 6:00-7:00 p.m; the room will be listed in the conference program.

Announcements

APSA Executive Commitee Nominations (2005-2007)

Treasurer:

Karen Remmer, Duke University. Professor Remmer's research has explored the politics of democratization and macroeconomic policy choice in Latin America.

At-Large Members of the Executive Committee:

Jeffrey Herbst, Princeton University
Professor Herbst's fields of interest include African politics and relations between industrialized and developing countries.

Mathew Shugart, University of California at San Diego
Professor Shugart specializes in constitutional design and electoral rules.

Elections will be held at the annual business meeting of the Section. Alternative nominations may be made at that meeting or by petition to the President. Petitions require the support of at least five section members.

Nominations Committee: Mark Beissinger, chair (University of Wisconsin), Ruth Collier (University of California, Berkeley), Catherine Boone (University of Texas, Austin), Richard Samuels (MIT), James Fearon (Stanford University), and Russell Dalton (University of California, Irvine).

In this Newsletter:

Volume 15, Issue 2
Summer 2004

How to Subscribe

Subscription to the APSA-CP Newsletter is a benefit to members of the Organized Section in Comparative Politics of the American Political Science Association. To join the section, check the appropriate box when joining the APSA or renewing your Association membership. Section dues are currently $8 annually, with a $2 surcharge for foreign addresses. The printing and mailing of the Newsletter is paid for out of member dues. To join APSA, contact:

American Political Science Association
1527 New Hampshire Ave., NW
Washington, DC 20036, USA

Telephone: (202) 483-2512
Facsimile: (202) 483-2657
e.mail: membership@apsa.com

Changes of address for the Newsletter take place automatically when members change their address with the APSA. Please do not send change-of-address information to the Newsletter.

 Published with financial assistance from the College of Arts and Letters, the Helen Kellogg Institute for International Studies, the Nanovic Institute for European Studies, and the Graduate School of Arts and Sciences at the University of Notre Dame du Lac.

APSA-CP Newsletter
University of Notre Dame
Decio Hall, Box “D”
Notre Dame, IN 46556
USA

Non Profit Organization
U.S. Postage
Paid