

Strategies of Judicial Research: Interviewing U.S. Supreme Court
Justices and Interest Group Attorneys

Lee Epstein

Southern Methodist University

Paper prepared for delivery at the 1986 annual meeting of the Southern Political
Science Association, Atlanta, Georgia

At a recent meeting of the Midwest Political Science Association, a curious event took place during back-to-back judicial panels. At the first panel, three scholars presented papers on the decisional outputs of individual Supreme Court Justices. Two of the papers took predominately "traditional," jurisprudential approaches to decision-making. One, however, attempted to build a model incorporating fact patterns to predict Justice White's voting behavior in search and seizure cases. The discussant, Harry Hirsch, applauded the efforts of the first two scholars, but lambasted the model-builder. Professor Hirsch was appalled by the "crass empiricism" of this scholars' work and directed him to more traditional, "softer" sources including biographies, papers, and possibly interview data.

During the next panel, focusing on access to the Supreme Court, once again three scholars presented papers. Two of the papers used "quantitative" analytic techniques to assess access to the Court, while the third was based solely on interviews with Justices, clerks, and other Supreme Court personnel. The discussant on this panel, Harold Spaeth of the Michigan State University, reacted quite conversely to Professor Hirsch: Professor Spaeth kindly critiqued the quantitative efforts, while severely chastising the paper based on soft data. He went so far as to call the research "smacking of insider-dopisms" and essentially useless in the absence of empirical verification.

This rather small incident illuminates what may be a growing and tumultuous debate within the judicial subfield, a debate centering around the respective roles of "hard" versus "soft" data.<1> This paper, of course, does not seek to end this discourse. Rather, based on my experiences both with "hard" and "soft" data, I seek to add another voice to the larger debate.<2> And, in fact, given that Professor Spaeth's view is perhaps more well-accepted than that

voiced by Professor Hirsch, this effort is slightly unbalanced, dealing more heavily with the uses of soft data, in the guise of intensive interviews.<3> More specifically, in this paper I describe the possible uses of interview data with an eye toward moving the debate from the abstract to the concrete.

The Uses of Intensive Interview Data

The debate surrounding the use of elite interviewing and its place in the scientific study of the judicial process had not been particularly well articulated by either side. In the abstract, those opposed to the use of intensive interviews argue that they may be "unreliable"<4>, that they do not produce generalizable results<5>, and that they force the reader to place too much faith in the interviewer.<6> Although those who use intensive interviews may find it easy to rebut these charges<7>, this is not my intent. Rather, a more important endeavor, it seems to me, is to ignore momentarily the abstract debate and address a question often relegated by both sides to a lesser status: Why bother interviewing elites? More specifically, what are the possible uses of intensive interviews?

To address this question, I have drawn on the works of other scholars, primarily writing on topics of the philosophy of science and research design. A general consensus exists among these researchers that intensive, elite, unstandardized, unscheduled interviews have three general purposes: exploratory, main research instruments, and supplementary (see generally, Kerlinger, 1973).

Exploratory Uses

When scholars discuss "exploratory research," they generally refer to a study that "proceeds without hypotheses and/or involves looking at a smaller amount or a different kind of data than would be appropriate for research designed to test hypotheses" (Bernstein and Dyer, 1984). Exploratory research, in short, is a legitimate way by which scholars develop hypotheses for future analysis.

Intensive interviews often provide an excellent method by which to collect exploratory or preliminary data. That is, often interviewees will suggest, perhaps "accidentally," new ideas for research that scholars can then translate into hypotheses for future testing. As Kerlinger notes, an interview "can be an exploratory device to help identify variables and relations, and to guide other phases of the research" (1973, p.480).

This is not a unique or strange occurrence. Quite often during the course of interviews elites make statements that they hold to be "fact," but which scholars have not considered in any serious way. Consider this example: While interviewing five Supreme Court Justices about the quality of group representation before the Court, several mentioned state attorneys. These Justices noted that some states had clear advantages (or disadvantages) in litigation because the quality of their representation was so varied, ranging from superb to intolerable. Although the Justices provided different and mostly vague explanations for the existence of such variation, their comments in and of themselves led to several interesting research questions: do states, in fact, have varied success rates before the Court and if variation does exist, can we systematically explain it?<8>

Another example of the utility of elite interviewing as an exploratory device came to us quite by accident. While interviewing representatives from conservative litigating groups, we noted a common denominator among those to whom we had spoken-- almost all mentioned that they had been involved with Barry Goldwater's 1964 quest for the presidency. Some had served as major consultants to the campaign, while others went door-to-door, handing out campaign literature. Even with such varied activity, we were astounded by the linkage among these conservative elites. Once again, this discovery led to some interesting research questions: did the rise of liberal public interest law parallel the conservative movement, for how long was such a trend evident, and is this a new variation on the old revolving door phenomenon?<9>

Hence, one important purpose of elite interviewing is exploratory in nature-- elites often provide excellent suggestions for innovative research hypotheses. Yet, as is the case with other sorts of exploratory studies, the researcher must act cautiously because elites are often sources of tremendous bias. As Manheim and Rich have indicated, the utility of intensive interviewing as an exploratory device may be negated if respondents "1)have so narrow a view of the events in question that they do not understand which aspects are important in explaining them; 2)have inaccurate information...; 3)have convinced themselves, in order to rationalize their own actions that things are one way when in fact they are another; or 4)intentionally lie in order to protect themselves or others" (1986, p.133).

In short, never substitute novel research suggestions for fact. Statements made by elites may be suggestive, may even seem "on their face" to be valid, yet are often tainted by perspective. In fact, I can cite as many examples of "suggestive research ideas" made by elites that, upon empirical analysis, proved

to be as worthless as the examples used above proved worthy. Consider this illustration: One week after a pro-life group, Americans United for Life, argued a 1983 case involving the constitutionality of local restrictions on abortion procedures¹⁰, we interviewed AUL attorneys. In passing, they mentioned that the organization may not win the case, but that they were sure that had Justice Steven's vote. Why? Because he seemed sympathetic in his lengthy questioning during oral argument. While looking over my notes on the interview, I thought that such an observation about oral argument could be transformed into a testable hypothesis. I gave up the idea several weeks later when the Court produced a strongly worded majority opinion upholding the principles enunciated in Roe v. Wade (1973) on to which the "sympathetic" Justice Stevens signed.

In sum, respondents' answers during intensive interviews provide a unique and often valuable source of preliminary data. As is the case with other research tools designed to gather data for the preliminary formulation of hypotheses, the researcher must proceed cautiously, never substituting novel research ideas for facts.

Interview Data as a Main Research Tool

The vast majority of scholars examining judicial voting behavior rely heavily on Supreme Court opinions (located in one of four reporter systems) as "the" data source. When we move beyond many institutional dimensions of the Court, however, we also must look to other sources of information. This is especially true for those studying interest group litigation. Most research in this area has, in fact, followed in the tradition established by the pioneering works of Clement E. Vose-- to "understand" a group, scholars must go out and

interview organizational representatives and staff attorneys (see Vose, 1959, 1972). From what other sources could we most accurately answer questions including how did the group form, why did it turn to litigation, what strategies does it use, and what impact does it believe it has? When scholars of the group process (or of other areas) use interviews in this way, "the products of interviews, respondents' answers to carefully contrived questions, can be translated into measures of variables...[and] are then to be considered as items in a measurement instrument, rather than as mere information-gathering techniques" (Kerlinger, 1973, p.480).

In this spirit, I too have used interviews as primary measurement instruments. Consider this research example: One of the most interesting conservative public interest law firms is the Mountain States Legal Foundation (MSLF). Located in Colorado, this organization was one of several formed in the mid-1970s to restore an ideological balance to the legal system. This firm is particularly interesting because 1)it is funded primarily by the Coors Foundation and 2)its first president was former Secretary of the Interior, James Watt. This much I discovered from interviews with others involved in the conservative movement. But many questions relevant to my study remained unanswered: what were the sepecific circumstances surrounding the founding of MSLF, how did Watt and Coors get involved, etc. As far as I could tell, only two sources could answer such questions, Coors-- which refused to grant me an interview and James Watt-- who did. Watt, while serving as Secretary of the Interior, in fact, gave generously of his time, fully addressing my questions.

Because Watt's responses to my questions became my primary measure of the concepts I was attempting to operationalize, they needed to be subjected to the traditional measurement standards of reliability and validity. As scholars are

very familiar with these, no need exists here to delve into much detail. Given the rather peculiar nature of intensive interviewing (e.g. scholars ask readers to place a great deal of faith in them), however, let me make a few general comments about evaluating answers as measures. To ensure "reliable and valid" responses, intensive interviewers should follow some rather straight forward rules. First, even though these are largely "unscheduled" interviews, use appropriate techniques. Many new textbooks on data-collection procedures (and specialized books, see Dexter, 1970) now regularly include chapters or sections on intensive, unstandardized, and unscheduled interviewing (see Babbie, 1982,1986; Williamson et al., 1982). Manheim and Rich (1986), for example, provide a list of 15 rules to follow while interviewing. Their major point in enumerating these suggestions-- even though these sorts of interviews differ substantially from formal survey settings, certain general guidelines should be followed. Second, researchers must work at minimizing interviewer bias (see Nachmias and Nachmias, 1981). Third, know your subject. That is, it is very important to understand the perspective from which your respondent is coming. Doing so will help to nullify often-voiced complaints of elite interviewing such as "researchers do not know enough about their subjects to recognize 'incorrect statements' or to analyze perceptively 'responses for possible sources of invalidity'"(Manheim and Rich, 1986).

Perhaps the most important lesson I have learned about using respondents' answers as measurement devices, however, is that researchers should not rely on a single respondent. Just as "quantitative" scholars, as we now know, should not depend exclusively on a single court era to evaluate the utility of certain theories (see Ulmer, 1986) or on one term to "measure" a Justice's at-

titude (see Ulmer, 1979), intensive interviewers must get as much information as they can from as many sources as possible.

In the case of James Watt and MSLF, I talked to at least five other individuals who had some knowledge, however tangential, of MSLF's role in the conservative legal movement. Happily, their comments well-supported those made by Watt. In other instances, I have been far less fortunate. Consider another example emanating from my experiences with conservative public interest law. When I started to look into the conservative movement, I came upon the existence of the National Legal Center for the Public Interest (NLCPI). As far as I could tell from its materials, this Center appeared to be a major part of the movement: its brochures and reports stated that it had been established specifically to facilitate the creation of conservative public interest law firms throughout the United States (including MSLF) and, then to coordinate these firms' efforts. Naturally, I thought to myself, "wow, this Center is the heart of the conservative movement and I should start my investigation with it."

Sure enough, I went down to the Center's offices in Washington, D.C. and listened patiently while its leaders told me of the major role they play in coordinating conservative litigation throughout the country: in fact, according to these individuals the NLCPI was "the" hub of the movement.

During the course of the interview, I must admit I was a bit skeptical of the information I was being "fed." The office in which I sat did not appear to be a "working" office-- very few people were there; it was incredibly vacant of any papers, legal or otherwise; and the individuals with whom I spoke simply did not possess enough substantive information about the groups over which they claimed to control and coordinate.

Over the course of another year of interviewing, indeed, my initial suspicions were confirmed. The groups "involved" with the NLCPI claimed that the organization in theory played the role of coordinator, but in practice had almost no influence over "their" constituents. Some of NLCPI's "groups," in fact, were so disillusioned with the organization that they considered themselves to be completely independent! Hence, the NLCPI, which on paper looked to be a major factor in the whole conservative litigation game, turned out to be just that-- a paper lion. Naturally, if I had not bothered to verify the information they had given me, my conclusions would have been just as "good" as my measurement instrument, which in this instance was worthless.

In sum, when scholars use interviewees' answers as the main instruments of their research, that is when "questions are designed to measure the variables of research," they must subject their operational definitions to the same standards normally applied. Some of these are peculiar to intensive interviewing; others are merely replicas of traditionally-used approaches.

Supplementary Uses

Although I have used interview data in an exploratory fashion and as a main research instrument, its use as a supplementary device is most compelling. When scholars refer to the supplementary utility of intensive interviewing, they do not necessarily mean that interviews provide "additional" information. Rather, interviews can supply important tools for "validating" the validity of one's measures. In fact, I have found that intensive interviews provide one of the best ways to evaluate, via construct validations (see Carmines and Zeller, 1979; Babbie, 1986), the measurement quality of operational definitions. As Kerlinger

as noted, "the interview can supplement other methods: follow up unexpected results [and] validate other methods..." (1973, p.480).

Let me demonstrate this through a number of examples. Five years ago, we focused a research effort on employment discrimination litigation. Generally we were interested in determining whether mid-level theories of interest group litigation, developed in the 1950s, held true for the Burger Court era. One of our hypotheses was that interest groups would be more successful than non-organized parties. Operationally we defined success as the number of wins/number of appearances. And, indeed, we found that groups were more successful than their counterparts. But, contrary to our expectations, the NAACP Legal Defense Fund was not the most successful organization-- the AFL-CIO won 89 percent of its cases (see O'Connor and Epstein, 1982).

Given this puzzling finding, we began to question the validity of our measure-- was the proportion of success actually measuring organizational success? One way to validate this measure, we decided, was to ask Supreme Court Justices about "success"-- were certain groups "better" than others and, if so, why? Although we felt it inappropriate to discuss specific organizations with the Justices, at least two volunteered such information. And, much to our surprise, in the course of their discussions both mentioned the AFL-CIO as one of, if not, the best litigator before the Court.

In another case, we received post hoc validation of our operational definition-- group cooperation. Again, in the course of studying interest groups that litigate, we hypothesized that certain kinds of organizations would be more likely than others to cooperate in court cases. In a paper examining the role of women's rights groups in sex discrimination suits, we operationally defined support as the number of supportive cases/total cases entered. Based on this

formula, we accorded each organization intergroup support scores, ranging from .090 to .846 (see O'Connor and Epstein, 1983). Given our knowledge of the groups involved these scores seemed to provide an accurate reflection of the "true" relationships among the organizations--e.g. face validity. Our operational definition of support, however, was more fully validated when an attorney from one of the represented groups called, mentioning that our support scores indeed accurately depicted the relations among women's organizations.

Both of the above stories have "happy endings"-- that is, interviews validated empirical definitions. What if, however, empirical measures and interview data yield conflicting results? Once again, consider the question of group success. Another way we tried to validate our empirical definition was to talk to group attorneys including representatives from the NAACP LDF. During the course of my interview, I pulled out a list of the employment cases the group had lost and won. The attorney and I started to discuss each of these, but in doing so, he would not admit that the organization had ever lost a case. For every "apparently" lost suit (i.e. the LDF asked the Court to reverse and the Court affirmed), this attorney found a redeeming point-- a small win on principle, the collection of attorneys' fees, etc. To make matters worse, I have found that almost every organization approaches their Court loses in much the same way as the LDF: they claim small victories within mountains of loses. How does the researcher deal with this? That is, how should scholars proceed when interview data seem to "invalidate" empirical/operational definitions?

To some extent, researchers can deal with this dilemma in the same way that they would approach any multiple indicator problem (see Carmines and Zeller, 1979; Babbie, 1986). For those who use interview data, however, it becomes particularly important to "go back to the source of the data." Consider my ex-

perience with the LDF attorney; it does not require any extraordinary acumen to understand why he wanted me to view the organization as a complete success story. As Manheim and Rich have noted, in these sorts of instances the respondent's "knowledge" is colored by perspective. But must we then necessarily cast off the interviewee as an "invalid" source of information?

The answer, I think, is no. In the instance of group success, for example, I have tried to incorporate a group's perspective in some of the following ways. First, when operationally defining group success, I am now careful to specify precisely what such a definition entails-- be it the group's or the Court's. Second, and more typically, I present "success" as a concept with multiple operational definitions. In Conservatives in Court (1985), for example, I measure success via three operations: the group's work as a whole, the percentage of cases in which the group supported the victorious party, and the group's assessment of its activities. In some instances, these measures all pointed in the same direction; in others, I had to evaluate fully differing "signals" before I could reach any definitive conclusions. In any event, I would never have considered a multiple indicator approach to success had I not had that telling conversation with the LDF attorney!

In sum, interviews provide an extremely compelling device to validate operational measures. Regardless of whether interview data do or do not support empirical measures, they can help scholars to understand more fully the objects of their endeavors.

Although scholars have criticized the use of intensive interviewing as an unreliable, non-generalizable method by which to collect data, this essay has tried to indicate that "elite" responses can play an important role in the research process. When scholars take appropriate precautions, intensive interviews can serve as vehicles by which to explore research questions, supplement existing measures, and, at times, as the main instruments of research.

Notes

<1>Certainly, the above example may better juxtapose the views of the discussants more than the larger debate. Professor Spaeth is well-known for his numerous efforts to predict/explain Supreme Court voting behavior via the Court's reporting systems. Professor Hirsch, on the other hand, examined Justice Felix Frankfurter vis-a-vis softer data sources, including manuscript collections and oral histories (see Hirsch, 1981).

<2>Over the past several years most of my research has explored the world of interest group litigation. My data sources have included: the opinions of the U.S. Supreme Court (as reported in U.S. Reports) the BNA microfiche brief series (all briefs submitted to the Court by sponsors and amicus curiae); LEXIS (a legal information retrieval system); and, most relevant for this paper, intensive interviews with two categories of persons: representatives of organized interest groups and governmental officials. The former includes attorneys and staffers from a wide range of ideological and issue-oriented groups. Some of the governmental officials with whom I spoke were: five U.S. Supreme Court Justices, Edwin Meese, James Watt, and Michael Horowitz (of the Office of Management and Budget).

<3>The kinds of interviews to which I refer have the following characteristics: 1)They are sometimes called "elite interviews," but not because the interviewee is necessarily an elite. As Manheim and Rich explain, "Their elite status depends not on their role in society but on their access to information that can help answer a given research question..." (1986, p.133). 2)They are largely unstandardized as "each respondent is treated differently to the extent that obtaining the information he or she alone possesses requires unique treatment" (Manheim and Rich, 1986, p.132). 3)They are largely unscheduled, that is, "no standardized schedule of questions is used" (Williamson, et al., 1982, p.173).

<4>Here critics question whether interviews with the same set of elites over and over again (by the same or different researchers) will yield the same or similar responses. Certainly, this is a valid issue: elite responses to questions often constitute, in research endeavors, operational measures of some concept. If this is the case, then such responses must be judged by traditional measurement criteria.

<5>Naturally, critics argue that intensive interviews do not produce "generalizable" results. Complaints here range from classic time-bound arguments (e.g. interviews are inherently situational) to the "deviant" case idea (e.g. the individual(s) selected for analysis represents the anomaly). Once again, these kinds of complaints have some validity-- it is almost impossible to interview all the individuals involved in any given specific research setting and to do so longitudinally is extraordinarily difficult.

<6>Scholars claim that intensive interviews force readers to place too much faith in the interviewer. Readers are asked to believe that the interviewer as followed prescribed techniques, has accurately recorded responses, and has asked the "right" questions.

<7>Most of those who rely on intensive interviews are acutely aware of these and other criticisms. In terms of the reliability issue, intensive interviewers simply note 1)that if proper procedures are followed, interviews are just as "reliable" as other forms of data collection and 2)that reliability is a necessary, but insufficient indicator of good operational measures. Such indicators must also be "valid"-- measurement procedures must reflect "the properties and relations that [they] purport to reflect" (Smith et al., 1976, pp.107-08).

The generalizability charge is more easily rebutted: intensive interviewers only have to point to the work of those using quantitative approaches, which are equally as problematic. In this past issue of the American Political Science Review, in fact, S. Sidney Ulmer (1986) demonstrated that social background theories are time-bound. Perhaps the most difficult charge intensive interviewers must face is that of "faith." Here, scholars tend to point to the fact that a prescribed set of guidelines exist for intensive interviewing. If these are followed, scholars argue, then this sort of interviewing is just as systematic as other forms of data-collection.

<8>Both of these questions can be answered affirmatively. We found tremendous variation in state success rates before the Court, ranging from .25 to 1.00. We were able to explain such variation through three sets of factors: litigant characteristics, legal expertise, and appeal rates (see O'Connor and Epstein, 1985).

<9>In a preliminary effort, we began to address these questions. Indeed, we found that a similiar revolving door pattern among liberal groups, going back as far as the late 1800s.

<10> Akron v. Akron Center for Reproductive Health, 462 U.S. 416.

References

- Babbie, E. 1982. Social Research for Consumers. Belmont, California: Wadsworth.
- Babbie, E. 1986. The Practice of Social Research. Belmont, California: Wadsworth.
- Bernstein, R.A. and J.A. Dyer. 1984. An Introduction to Political Science Methods. Englewood Cliffs, New Jersey: Prentice-Hall.
- Carmines, E.G. and R.A. Zeller. 1979. Reliability and Validity Assessment. Sage University paper series on Quantitative Applications in the Social Sciences. Beverly Hills & London: Sage Publications.
- Epstein, L. 1985. Conservatives in Court. Knoxville, Tennessee: University of Tennessee Press.
- Hirsch, H. 1981. The Enigma of Felix Frankfurter. New York: Basic Books
- Kerlinger, F.N. 1973. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston.
- Manheim, J.B. and R.C. Rich. 1986. Empirical Political Analysis. New York: Longman.
- McGraw, D. and C. Watson. 1976. Political and Social Inquiry. New York: Wiley.
- Nachmias, D. and C. Nachmias. 1981. Research Methods in the Social Sciences. New York: St. Martin's.
- O'Connor, K. and L. Epstein. 1982. The importance of interest group involvement in employment discrimination litigation. Howard Law Review, 25:711-729.
- O'Connor, K. and L. Epstein. 1983. Beyond legislative lobbying: Women's rights groups and the Supreme Court. Judicature, 67:134-143.
- O'Connor, K. and L. Epstein. 1985. States as Litigants. Paper delivered at the annual meeting of the Southern Political Science Association, Nashville, Tennessee.
- Smith, B.L. 1976. Political Research Methods. Boston: Houghton Mifflin.
- Ulmer, S.S. 1979. Parabolic support of civil liberties claims: The case of William O. Douglas. Journal of Politics, 41:634-639.
- Ulmer, S.S. 1986. Are social background models time-bound? American Political Science Review, 80:957-967.

Vose, C.E. 1959. Caucasians Only. Berkeley, California: University of California Press.

Vose, C.E. 1972. Constitutional Change. Lexington, Massachusetts: Lexington Books.

Williamson, J.B., et al. 1982. The Research Craft. Boston: Little, Brown.