

HANDBOOK OF POLITICAL SCIENCE

Volume 7

***STRATEGIES
OF INQUIRY***

Edited by

FRED I. GREENSTEIN Princeton University

NELSON W. POLSBY University of California, Berkeley

© 1975



ADDISON-WESLEY PUBLISHING COMPANY

Reading, Massachusetts

Menlo Park, California • London • Amsterdam • Don Mills, Ontario • Sydney

SURVEY RESEARCH

RICHARD W. BOYD
with
HERBERT H. HYMAN

INTRODUCTION

There is no doubt that survey research has made a major contribution to the postwar growth of political science. In our more expansive and confident moments, we would agree with McClosky's assessment that "within two decades the survey method has become the most important procedure in the 'behavioral' study of politics . . ." (1967, p. 65). Now and again, however, we are uncomfortably reminded that no method is infallible. With great chagrin the July 1973 issue of the *Current Population Reports*' series, *Consumer Buying Indicators*, announced the end of the Survey of Consumer Buying Expectations. A special study of the survey had revealed the disconcerting fact that from 1967 through the first half of 1972 the correlation between the index of expected car purchases and actual new car purchases was negative. Moreover, the index had such a large standard error that there was no statistically significant change in the predictive index from July 1967 to April 1971, "a time of considerable variability in actual purchases," the report dutifully noted (*Current Population Reports*, Series P-65, 46:1). So we will not argue that the survey method has not had its pratfalls. We simply will not dwell long on them, proceeding instead on the more comfortable assumption that one learns more from examples of successes than failures.

In spite of or, more properly perhaps, because of the variety of surveys, we need to circumscribe our task. Should we begin with the classic nineteenth-century surveys of poverty in England by Booth and Rowntree, even though

We would like to thank Russell D. Murphy, Leon V. Sigal, and Clement E. Vose of Wesleyan University for their helpful suggestions of examples included in the chapter and Angus Campbell of the University of Michigan for reading the section on the fusion of research and policy interests in the development of survey techniques.

their work could not take advantage of sampling techniques yet to be invented? Should we include Lasswell's *Psychopathology and Politics* (1930), even though another researcher could not replicate his unstructured interviews? We prefer a more limited definition of a survey in order to eliminate the extraordinary number of studies that form a penumbra around the core group of studies possessing all the components of survey methods. We will restrict the definition of a survey to one formulated for a prior analysis of this literature: "A survey is an inquiry of a large number of people, selected by rigorous sampling, conducted in normal life settings by explicit, standardized procedures yielding quantitative measurements" (Hyman, 1973, p. 324).

As a technique of data gathering, survey research falls into the larger category of field methods. In their rigid requirements of sampling from strictly delimited populations and in their standardized procedures for acquiring information, surveys clearly depart from the more traditional methods of empirical political science. The memory of de Tocqueville, Martineau, and Bryce as exemplars of work based on wise and detached observation and reflection makes one wonder if the accoutrements of modern techniques are really necessary. On the other hand, we might also infer from these notables that a researcher who is not a genius had better bring along his bag of tools.

In addition to the observations of world travelers, surveys contrast also with other traditional methods of data collection. In all, the richest sources of political data have undoubtedly been documents and aggregate data produced by governments and private groups. V. O. Key, Jr., surely a most important political scientist, made most of his major contributions via shrewd analyses of aggregate data. But in his two final books, *Public Opinion and American Democracy* (1964) and *The Responsible Electorate* (1966), he demonstrated how important it is to have data about people as individuals in order to make statements about people as a corporate entity.

Key did not, as E. E. Schattschneider once said of public opinion analysts, confuse the actors with the audience. Key knew that the initiatives for policy change and the selection among alternative programs usually rests even in a democracy with governmental leaders. But his argument in *The Responsible Electorate* (1966) focused political scientists' attention on voters' evaluations of government programs, particularly during what he called "critical" (realigning) elections. His timely argument encouraged researchers to look for key issues as the bases of voter decisions in the United States elections of the 1960s and 1970s. (See, for example, the articles and citations in the June 1972 *American Political Science Review* symposium on issue voting.) It is difficult to imagine how Key could have ever made a credible argument that the electorate is indeed rational without having at his disposal the data on individual voters' opinions that survey research provided.

Similarly, Robert Dahl's first works relied in the main on the traditional

sources of data in political science: elite interviewing, government documents, and the writings of major philosophers. To be sure, *Congress and Foreign Policy* (1950), in its analysis of the dilemma of citizen competence and democratic control of foreign policy, did utilize secondary analyses of opinion polls of the 1930s and 1940s. However, Dahl's principal concerns focused on party cohesion and responsibility in Congress on foreign policy issues, for which he relied mainly on roll call votes and interviews with critical members of Congress and its staff. In *Politics, Economics, and Welfare* (with Charles E. Lindblom in 1953) and in *A Preface to Democratic Theory* (1956), Dahl made landmark contributions to the problem he had earlier posed concerning the relationship between citizens and leaders in a democracy. This he termed the "First Problem of Politics": How to prevent leaders from becoming tyrants (1953, p. 273). Dahl noted the importance of training leaders in democratic norms; however, his greater stress was on the pluralistic organization of social and political power, without which "no constitutional arrangements can produce a non-tyrannical republic" (1956, p. 83).

Though Dahl became with these books the leading interpreter of the pluralist school, his work still lacked direct evidence on the character of power relationships between citizens and political leaders. That lack is presumably one reason why he titled his treatise on democratic theory, *A Preface . . .* In the late 1950s, he filled this lacuna with coordinated surveys of the people in New Haven, Connecticut: personal interviews with important local decision makers; a questionnaire sent to all members of a well-defined population of subleaders in political parties, a redevelopment agency, and Board of Education and PTA officials; and finally personal interviews with two samples of registered voters. From this design came empirical answers to the issues he had raised. In *Who Governs? Democracy and Power in an American City* (1961), he detailed a system of power in which there existed a different set of leaders and attentive publics for distinct domains of policy. It was the evidence for pluralism that he needed to support the arguments of *The Federalist Papers* he had analyzed in *A Preface to Democratic Theory* (1956).

What then do we infer from the patterns of research of Key and Dahl? It is certainly not that survey research has supplanted documentary and aggregate evidence as the most important data source in political science. Rather, we have drawn attention to the work of two scholars who made profitable use of multiple data sources and who both came to view information on the policy opinions and democratic values of individuals as critical to lines of inquiry linking, in both cases, several books. This chapter, then, is no brief that empirical research questions can only be answered with surveys; it does suggest, however, that only surveys can answer *some* questions. (Another pointed example is Wilson and Banfield's (1971) use of a survey in an attempt to resolve a controversy over the "political ethos" of immigrant and Anglo-Saxon subcul-

tures.) Beyond this point we emphasize the quite extraordinary range of valid research techniques and the points at which surveys can help settle crucial elements in a broad gauged theoretical investigation.

To summarize, surveys provide data on a large number of people in their varied and normal life settings. They provide quantifiable measures on a notable range of behaviors and beliefs, thus opening to the researcher the analytic powers of an increasingly sophisticated statistics. And they do so using scientific bases for sample selection that justify confidence about the broader universe to which findings can be generalized.

Looking ahead, it is not our purpose to present a minitext on survey design and analysis. Our bibliography lists numerous examples of both general and specialized works on survey research methods. Instead, we will ask the questions that someone examining the potential of survey methods for the first time might ask: What kinds of belief and behavior can be studied? For what purposes? With what effect on substantive theory? Our goal is to suggest the ways in which surveys can answer particular questions by pointing to examples in the substantive literature of political science, though in our search for apt illustrations we will occasionally stray outside the domains of politics and policy analysis.

WHAT KINDS OF BELIEF AND BEHAVIOR CAN BE STUDIED?

Demographic Attributes and the Social Survey Movement

A primary use of surveys is in demography. Indeed, the origin of modern survey research lies in the social and health survey movement of the nineteenth century. Parten (1950, p. 6) credits Charles Booth with conducting the first comprehensive social survey in London in 1886, published as *Life and Labour of the People of London* in 17 volumes (1892–1897). This survey of economic and social conditions was designed to show, as Parten quotes Booth, “the numerical relation which poverty, misery, and depravity bear to the regular earnings and comparative comfort and to describe the general conditions under which each class lives” (Parten, 1950, p. 6). B. Seebohm Rowntree followed Booth with an investigation of the conditions of the working class in York, England: *Poverty, A Study of Town Life*, first published in 1901.

The Discovery of Poverty in the United States, as Bremner so neatly subtitles his book *From the Depths* (1956), provides a good history of the development of America’s counterpart to the European social survey movement. Complementing portraits on slum life in the nineteenth century by popular writers such as Jacob Riis were Carrol D. Wrights’ survey of unemployment in Massachusetts during the depression of 1873–1879 and the Association for Improving the Condition of the Poor’s Survey of New York City in 1873 (Stephan, 1948,

p. 18). Similarly, the Citizens Association of New York sponsored the *Report of the Council of Hygiene and Public Health* (1865), an extensive investigation of sanitation facilities and disease in New York.

The question of whether or not we would label these early demographic investigations as surveys is ambiguous. Several meet our criterion of standardized procedures of measurement. For example, the latter report on New York City includes a detailed list of instructions to investigators on the delimitation of the area surveyed, methods of inspecting houses, and queries of household residents on sickness and mortality (pp. xxiv-xxxv). These nineteenth-century social surveys fail to meet our criteria for surveys only for their lack of scientific sampling. As Stephan notes (1948, p. 21), "although the theory of probability was well established in the eighteenth century, its applications to the practical drawing of samples were delayed until the twentieth." Not until 1912, when Arthur Bowley surveyed the economic conditions of the working class of Reading, England, were the principles of scientific sampling first used. Moreover, even after Bowley's sampling innovations had been published, United States surveys still followed Booth's procedures of attempting to catalogue complete populations of cities on a broad range of subjects (Parten, 1950, p. 9).

The social survey movement illustrates one rather common feature of surveys of social and demographic data: the interpretation of the data is typically grounded in politics and policy preferences.

First of all, the motive of the social surveyors was to dramatize the problems of the poor in order to affect policy in areas of health, sanitation, living conditions, and working conditions. The perspectives of the investigators, of course, varied. Some saw the problems from the point of view of the poor and hoped to stimulate substantial changes in their lives. The fruits of their efforts were realized in such social legislation as the New York Tenement Inspection Act of 1901, the first effective inspection law and one that became a model for other cities (Beard, 1912, ch. 11; Davies, 1966, p. 4).

Others in the social survey movement seemed more concerned with the threat to social order posed by the living conditions of the poor than by a regard for the welfare of the poor as such. For example, the introduction to the survey report of the previously mentioned Council on Hygiene and Public Health commented:

But beyond the physical, the mental, and the economical losses resulting from prevailing ill-health, there are certain political and social aspects of the same agencies that ought to be studied by every intelligent citizen. The mobs that held fearful sway in our city during the memorable outbreak of violence in the month of July, 1863 [the Civil War draft riots], were gathered in the overcrowded and neglected quarters of the city. As was stated by a leading journalist at that time: "The high brick blocks and closely-packed houses where the mobs originated seemed to be literally *hives of sickness*

and vice. It was wonderful to see, and difficult to believe, that so much misery, disease, and wretchedness can be huddled together and hidden by high walls, unvisited and unthought of, so near our own abodes. . . . Alas! human faces look so hideous with hope and self-respect all gone! And female forms and features are made so frightful by sin, squalor, and debasement! To walk the streets as we walked them, in those hours of conflagration and riot, was like witnessing the day of judgment, with every wicked thing revealed, every sin and sorrow blazingly glared upon, every hidden abomination laid before hell's expectant fire. The elements of popular discord are gathered in those wretchedly-constructed tenant-houses, where poverty, disease, and crime find an abode." (Citizens Association of New York, 1865, pp. xv-xvi)

In a similar vein, Diamond (1963) details the obvious political motives that underlay many nineteenth-century questionnaires, including one example of a questionnaire taken among plantation owners designed to show that religious instruction did not lead slaves to subversive thoughts.

A second way in which the gathering of social and demographic data is connected to politics is that the type of data collected is a guide to what people view as the major social and political problems of a particular period of history. In the second half of the nineteenth century the problems of the poor were interpreted as the problems of cities, particularly problems linked to the passage of immigrants into city slums, to the political corruption that middle-class reformers saw immigrants as creating, and to the challenge of the then novel experiment with universal manhood suffrage. One can see these concerns in the dates that federal data first became available on a *city-by-city* basis: nativity (1850); occupation and education (1870); blind, deaf-mutes, and insane and feeble-minded (1880); citizenship and inability to speak English (1890); and housing (1900) (U.S. National Resources Commission, 1939).

Third, just as policy concerns cause data to be collected, existing data, however inaccurate, will usually be used to justify policy arguments. For example, Regan (1973) has recently discussed the deficiencies of the first six United States censuses of 1790 through 1840. Interest in reforming existing census procedures was peaked by the use that proslavery forces made of the 1840 figures on the proportion of "insane and idiots" among free and slave Negroes. The figures suggested that the proportion of Negroes enumerated in the North as "insane and idiots" was substantially higher than the proportion indicated for the South. Regan quotes Calhoun's use of the data, "Here is proof of the necessity of slavery. . . . The African is incapable of self-care and sinks into lunacy under the burden of freedom. It is a mercy to him to give him this guardianship and protection from mental death" (Regan, 1973, p. 541).

The figure for "insane and idiots" among Negroes was just one error of many produced by poor administration and non-standard interview schedules.

A young physician named Edward Jarvis campaigned for reform of census administration, and succeeded in persuading the American Statistical Association to take up the cause. As a result Congress adopted substantial changes in the law in preparation for the 1850 census. Because of the advances made in the administration of the 1850 census, "it marked the beginning," as the Director of the Census said in 1907, "of scientific census inquiry in the United States" (Regan, 1973, pp. 544-545).

Once the 1850 census had begun the process of standardizing interview schedules and adopting uniform procedures for organizing and displaying the data for all the states, the utility of the census for demographic measurements was assured. It only remained for the 1940 census to institute sampling in the United States census, in this case a 5 percent sample of households who were asked a list of supplemental questions in addition to those posed to the remainder of households, including the critical unemployment and income questions that have only been asked of census samples. As Parten notes (1950, p. 14), the adoption of sampling by the United States census blurs a distinction commonly made between a census and other surveys—that a census is a total enumeration, whereas most surveys are samples.

The policy impact of the census is, of course, enormous. United States federal grants-in-aid payments to states and cities are directly tied to census estimates of population, and the systematic underrepresentation of mobile, young members of minority groups often costs cities substantial federal funds to which an accurate census would entitle them. In terms of sample size, measurement error, and careful administration, however, the United States census has no peers among academic surveys. Its limitations are those imposed by the political sensibilities of a government that does not want to risk tainting the generally high reputation of the census with charges that it requests politically sensitive information. For example, the census asks neither questions of political attitudes nor of voting choices, though it does risk asking its sample respondents whether they are registered to vote and if they went to the polls. From these samples the Bureau has published an exceedingly useful series of publications on registration and voting patterns among social groups, the P-20 series of *Current Population Reports*.

Contemporary uses of surveys to measure demographic and social attributes are as intriguing as those of the social survey movement. One such example is the study of social and occupational mobility and its effects on political attitudes. Inter- and intra-generational mobility are two of the fundamental concepts in political sociology, since observing the impact of upward or downward movement in status highlights the influence of status relationships on political attitudes. One recurring question is whether or not mobility causes people to abandon the values they formerly held in favor of the values of the class or group into which they have moved.

The effects of mobility on values are nearly always more complex, even

intractable, than they first appear. Citing an example of this, Lipset and Bendix suggest that mobility may lead to conservatism in one situation and radicalism in another. Applying this observation to South Africa, they observe:

... if a Negro in South Africa obtains a nonmanual position, he is a ready candidate for leadership in a movement of radical protest. But if a White American from a working-class family makes the same move, he usually becomes politically and socially conservative. (Lipset and Bendix, 1964, p. 64)

Lipset and Bendix attribute the difference in the attitude changes of the hypothetical South African and American to a concept that has enjoyed a considerable vogue in recent years. Called variously status inconsistency, status discrepancy, and low status crystallization, this concept derives from the fact that we are all embedded simultaneously in a multidimensional status structure in which a person's position in the eyes of others may be determined by many of his statuses: his ethnic background, income, education, and occupation. Because a person may move up and down his individual status ladders at different rates, he may at once occupy a high position on one status and a low position on another.

Rush (1967) notes that the concept of status inconsistency has been proffered as an explanation of a range of social and political attitudes: a sense of social isolation, desire for fundamental radical change, motivations to action, conservatism, and political liberalism. He, along with others (Bell, 1963), uses the concept to explain right-wing conservatism.

Despite the disparity of the dependent variables to which it has been related, status inconsistency retains its conceptual focus. Most investigators assert that variation in their chosen dependent variable results from the psychological stress that inconsistent statuses produce. This stress is itself the product of frustration and uncertainty that grow in those who occupy inconsistent status positions. Inconsistency creates ambiguous expectations for the particular person as well as for those with whom he interacts. How is the black professional to be treated? How indeed does he see himself? As this example suggests, status inconsistency is usually thought to be particularly stressful for people who have earned a high position on an achieved status, yet occupy a lower position on an ascribed status such as race (Jackson, 1962; Jackson and Burke, 1965). Thus, psychological stress intervenes between status inconsistency and the political attitudes that the stress is presumed to influence.

The theoretical appeal of the concept of status inconsistency is manifest in a number of studies that have made it a central explanatory concept (Lenski, 1954; Goffman, 1957; Kelly and Chambliss, 1966; Treiman, 1966). For our purposes the point is that the concept could never be investigated empirically without survey research. It is not enough to have aggregate statistics on a

nation's racial, occupational, or income distributions. The concept requires a look at each individual's status *set*. What are his race *and* his occupation? Only when the individual is the unit of analysis can we investigate such a concept.

The American Occupational Structure, that classic of survey research by Blau and Duncan (1967), provides a second example of the political uses of demographic information in surveys. In a cooperative venture with the United States Census Bureau in 1962, they added a series of questions on inter- and intra-generational mobility to the Bureau's continuing *Current Population Survey*. The sample consisted of about 20,700 males between the ages of 20 and 64. Because the Census Bureau shuns questions on political attitudes for the understandable reason that it does not want to be charged with maintaining a government data bank on the political values of individual citizens, no political attitude items were included on the interview schedule. Nonetheless, Blau and Duncan devote much of their concluding chapter to speculation on the relationship of opportunities for social mobility in a society to its politics—its equality of political influence, its political stability, its capacity for change.

Much of their argument is a critique of Lipset and Bendix's view on the relationship of mobility to egalitarian values and stable democracy. Lipset and Bendix had challenged de Tocqueville's thesis that the opportunity for social mobility in the United States is the source of its political stability. This explanation fails, they say, because rates of movement between blue-collar and white-collar classes are no higher in the United States than in other industrial countries. They suggest instead a variant of Hartz's thesis in *The Liberal Tradition in America* (1955): that the United States possesses an egalitarian value system because it had no feudal past and no conservative aristocracy, and this is responsible for its political stability.

Blau and Duncan find this explanation unpersuasive and question what perpetuates this egalitarian value system if, in fact, rates of mobility in the United States are not higher than those of other countries. Instead, they point to the implications of a finding that Lipset and Bendix, with their more limited data base, could not have known: ". . . the opportunities of men originating in lower social strata to move into *top* positions in the occupational hierarchy are greater in the United States than in other countries," and this sustains its egalitarian values (Blau and Duncan, 1967, p. 437, italics added). Blau and Duncan also note that the "crass materialism of American values" elevates "possessions into the most important distinguishing feature of differential status," which, they argue, "makes it easier to translate economic improvements into advancements in accepted social status" (1967, p. 437).

Bottomore (1968) in turn criticizes this explanation by Blau and Duncan, noting periods in American history when both American intellectual and popular movements challenged as myth the assumption of an egalitarian society. He finds the explanation of political stability in the failure of protest

movements to establish themselves as an effective opposition party, a result, he argues, of the ethnic diversity of the population, the "peculiar" situation of Negroes, and ultimately the steadily rising standard of living.

However this argument is resolved, it underscores the importance of theories of social mobility and political values. And to risk the lesson in repetitive argument, one cannot measure mobility, not to mention political attitudes, without the measures of individual characteristics that survey research provide.

Perhaps the most convincing evidence of the usefulness of survey research in demography is the number of projects around the world that are replicating the Blau and Duncan study. These surveys comprise what Featherman and Sewell (1974) have termed a second generation of mobility studies, and the list of nations includes, in addition to their own study of the United States, Australia, Japan, France, Germany, Hungary, Poland, Canada, Denmark, Finland, the United Kingdom, Yugoslavia, Switzerland, and Italy. The Centre for Intergroup Studies of the University of Cape Town has begun a similar study of social mobility among the Coloured population group in the western Cape region of South Africa and plans to extend the study to include Africans and whites as well. In short, studies of social mobility and political values presently constitute an extraordinarily fertile field of research, all directly traceable to the special attributes of the survey technique: the possibility of measuring the characteristics of the individual efficiently and economically.

Attitudes, Beliefs, and Behavior

Attitudes, beliefs, and behavior constitute a second major category of social and political data that can be efficiently studied with surveys. However, simply to use these two words, attitude and belief, is to raise a controversy concerning the definitions of the terms. One group of attitude theorists regards the concept of attitude as a unidimensional concept. For example, Fishbein defines an attitude as pure affect for or against some object, and provides separate operational definitions for beliefs and behavioral intention. In this approach he follows in the steps of others such as Thurstone and Osgood. Other theorists, including Rosenberg and McGuire, regard attitude as a multidimensional concept, with affective, cognitive, and conative components.

The dispute over the more useful definitions of attitude is discussed in articles by Zajonc (1968), Scott (1968), and McGuire (1969) in the companion series to this handbook, *The Handbook of Social Psychology*. Since our task here is to underscore the variety of significant applications of surveys, there is no reason for us to take a position on the proper conception of attitude. Rather, we will be content to note, as Greenstein (1973) said about personality theory, that attitude theory is a pluralistic universe. We will group responses to surveys in various content domains (affect, attitudes and beliefs, judgments, behavior,

ideologies and belief systems, and cultural values) and emphasize in the discussion that few surveys are limited to a single domain. Quite the contrary, most surveys explore a constellation of responses across a rich and variable array of subjects, a fact that will be evident in the range of topics we discuss.

Political affect. To search the literature on political affect is to confront a paradox. Surveys of political attitudes and actions are the substance of endless surveys, with beliefs on issues, candidates, and parties repetitively measured with items that catch the subtlety and complexity of people's evaluations of politics. In contrast, surveys of affect, those "emotions and sentiments that accompany political thought and behavior" (Hyman, 1973, p. 345), are quite rare. The paradox is all the more pointed when we reflect that so much of our literature on political behavior argues that many people's expressed opinions are marked by lack of affect, by an absence of feeling and intensity. But how can we know this, when we have invested so few resources in measuring the emotions that people bring to politics?

To be sure there are a few examples of studies of affect, but their presence is marked by the isolation in which they stand. Almond and Verba (1963, pp. 145-146) asked the respondents of their five-nation survey to report whether they found "election campaigns pleasant and enjoyable," whether they "sometimes get angry during campaigns," whether they "sometimes find campaigns silly or ridiculous." The answers suggested substantially different emotional responses to campaigns between countries. For example, the majority of United States citizens admitted that they found them at times pleasant, at others ridiculous. At least one gathers that they found them emotional experiences on some level. In contrast, a majority of Italians reported that they "never enjoy, never get angry, and never feel contempt" during campaigns. These pointed differences in national responses suggest the value of studying political emotions and sentiments, and it is unfortunate that we have not pursued such suggestive research.

There does exist a pair of studies that have focused on measuring affect. In the first study Fishbein and Coombs (1971) sampled an unspecified mid-western American city using a panel design with waves of interviews before and after the 1964 election. In Fishbein's attitude theory, attitudes are defined precisely as "the amount of affect for or against a psychological object" (Fishbein and Coombs, 1971, p. 5). In his terms one can measure an attitude (e.g., toward government medical insurance) with evaluative scales using "good" and "bad" as their anchor points. A "belief" in turn is a cognitive judgment defined as "the probability or improbability that a particular relationship exists between the object of belief and some other object" (1971, p. 6). An example of the two measures will clarify the distinction between attitude (or affect) and belief.

1. An evaluation scale of *attitude (affect)* toward Medicare.
 Medicare

Good— : — : — : — : — : — : —Bad

2. A cognitive scale of *belief* about a candidate's position on Medicare.
 Lyndon Johnson is in favor of Medicare.

Probable— : — : — : — : — : — : —Improbable

In Fishbein's theory the behavioral intention to vote for Lyndon Johnson would be expected to be a function of a person's attitudes (affect) toward some issues such as Medicare and his beliefs about the candidate's views on the issues. As Fishbein and Coombs predicted, there was a high correlation in this sample between the reported vote in the postelection interview and vote predictions made on the basis of attitudes and beliefs toward a series of issue objects measured in the preelection interviews.

The second study by McClure and Patterson in Syracuse, New York is similar to that of Fishbein and Coombs in design, though not in theoretical focus (McClure and Patterson, 1973; Patterson and McClure, 1973; and Patterson, McClure, and Meier, 1974). Patterson and McClure (1973) were concerned with the effects of media, especially paid television advertising, in altering people's attitudes and beliefs about issues and candidates. As in the Fishbein-Coombs study, attitudes were defined as affect measured with evaluative scales, with "good" and "bad" as the polar anchor points. Their findings make important contributions to the conventional wisdom of mass communications theory. Media advertising did not change the viewer's attitude (or affect) toward issues such as defense spending or pulling troops out of Vietnam. But media did change people's *beliefs* about where the candidates stood on those issues, in the main contributing to voters' acquisition of accurate information about candidate positions on those issues.

As interesting as the Fishbein and Coombs, and McClure and Patterson, studies are, they do not alter our judgment that the study of political affect is sadly neglected. In the first place, by defining attitudes toward public policy issues in terms of "good-bad" evaluative scales, a rather strong cognitive component infuses the measures of affect. After all, we likely think that Medicare is good or bad because of other cognitive beliefs about the state of public and private health care facilities and people's ability to pay for them. Thus, affect in this usage is more cognitively based than the kinds of emotions and sentiments that we usually denote as affect.

In the second place, in these two studies affect was measured in only one of its manifestations, the evaluation of good and bad. Surely this is justifiable in terms of the length of time required to administer semantic differential scales. It does mean, however, that *qualitatively* different affective responses are ob-

scured. Ponder the variety of emotions and sentiments that sustain people in politics or repel them from it: anger, disgust, revulsion, fear, alienation, apathy, or the opposing sentiments of satisfaction, gratification, fulfillment. We rarely measure these emotions directly, even though most of us presume they are motive forces in our political behavior. Rather, we tend to infer the emotions from the cognitive content of people's opinions.

Consider, for example, the literature on those venerable political scales of the United States voting studies—the indexes of political efficacy and political cynicism. Taken singly, both bear consistent relationships to the act of voting. Feelings of low political efficacy typify those who rarely vote. Political cynicism, a mistrust of officials in office, is related to a disposition to vote against the candidates of the party in office in national elections (and to vote against bond and policy referenda in local elections) (Campbell *et al.*, 1960; Boyd, 1967; Gamson, 1968).

Taken together, low political efficacy and political cynicism are often considered as the cognitive and the affective components of political alienation. (See Table 1.) That is, an alienated person not only believes that he is powerless, he also resents the fact that he is, an inference from his cynicism and mistrust of public officials (Gamson, 1968). In fact, the alienated voter is but one of a number we can define using the extremes of the efficacy and mistrust indexes (Boyd, 1967).

TABLE 1 Combinations of political efficacy and political trust

		Trust of public officials		
		Low	Medium	High
Sense of political efficacy	Low	Alienated voter		Acquiescent voter
	Medium			
	High	Protest voter		Allegiant voter

Clearly, these four voter types are distinguished as much by their feelings or emotions about politics as by cognitive beliefs about their own ability to influence the course of politics through the voting process. The estrangement of the alienated voter has just been described: it is resentment grounded in felt impotence. This is a different attitude from that of the protest voter, who acts on his mistrust of political leaders because of his own convictions of political potency. The belief, as Gamson put it (1968, p. 48) that influence

is both "possible and necessary," has been one of the distinctive features of postwar American politics, establishing what Converse has called a "prime setting for the effective mobilization of discontent" (1972, p. 336).

We may think of cynicism as a direct measure of affect, but it is surely grounded in cognitive beliefs about the behavior of candidates in office. More accurately, we *infer* that emotional intensity must infuse people's political behavior, when particular combinations of feelings of political power and political mistrust interact. Thus, we have been content to measure the cognitive components of the attitudes directly and assume the affect that underlies them.

We have not, of course, exhausted treatments of affect in survey research. Sears, for example, in his review of the literature on political behavior for the *Handbook of Social Psychology* (1969), interprets the political socialization of children as the acquisition of partisan loyalties and allegiances toward public figures through the role of the family and schools in nurturing positive affect toward politics. Surely it is right that in the child's world, feelings about politics precede cognitive beliefs about the political system. Hyman (1959), Greenstein (1965), Hess and Torney (1967), Easton and Dennis (1969), and Jennings and Niemi (1974), as well as many others, have contributed to these theories of the development of affective loyalties toward the political system, exploring feelings of political legitimacy toward what they term the political community, the regime, and its authorities. In a series of articles, Arthur Miller has examined the breakdown of those affective allegiances in the United States and the growth of feelings of estrangement from politics (1973, 1974a, 1974b). Yet, when we reflect on the variety of qualitatively different emotional responses to politics, responses ranging from revulsion to adoration, we wonder why researchers have devoted so little of their energies to measuring those feelings that so strongly influence the character of our political behavior.

Political attitudes and beliefs. Measures of attitudes and beliefs comprise the bulk of most surveys. The usefulness of attitude surveys is so evident that we will not pause to make the argument. What may not be quite so obvious is the multidimensionality of this content domain, which we sense only when we ask ourselves what it is that we would like to know about a belief. Then the question focuses our attention on the complexity of what first might have appeared as a simple entity.

For example, we would ask the kinds of questions that Smith, Bruner, and White (1956), Lane and Sears (1964), and Best (1973) have asked: What is the *direction* of an opinion? Are the respondents in the survey in favor of more or less control of wages and prices? How much more? How much less?

How *intensely* does the person hold this view? With what level of conviction does he imbue this belief? Lane (1962) and Riesman (1950), for example, have argued that Americans rarely hold intense opinions, rarely look with outrage on those whose opinions are contrary or who convert their political

opinions into moral crusades. Freidson (1955) has emphasized the developmental processes that inhibit the formation of such extreme views. In a similar vein Almond, in his influential book *The American People and Foreign Policy* (1950), argued that the public is largely indifferent to most issues in foreign affairs:

The characteristic response to questions of foreign policy is one of indifference. A foreign policy crisis, short of the immediate threat of war may transform indifference to vague apprehension, to fatalism, to anger; but the reaction is still a mood. (Almond, 1950, p. 53)

This "mood theory" of foreign policy opinion is consistent with the arguments of many who assert that the public lacks the conviction and interest to support a sustained, costly role in foreign affairs. Still others argue on the same premise that the public will not likely support a long-term, conventional war, preferring instead either isolationism or extreme escalation of conflict (Key, 1964; Brzezinski and Huntington, 1964). Caspary (1970), among others (Boyd, 1972, p. 443), has raised questions concerning the mood theory. The controversy is an important one, and it underlines why we must know whether an opinion is crystallized, intensely held, and stable or whether it is weakly held, fluid, and easily changeable or manipulatable.

We would also ask about the attitudinal *structure* in which an opinion is embedded. Is the opinion part of a constellation of other beliefs? Is it organized into and supported by a more abstract political ideology? We will consider this issue in a later section on surveys designed to explore ideologies and belief systems, but we mention it here to note that the most important aspect of what we may want to know about one belief is its relationship to another.

We would want to ask as well about the *level of information* that supports a belief. Even at a minimal level many people lack information on major policy choices of their government. For example, in 1969 the United States Senate Foreign Relations Committee held televised hearings on President Nixon's proposal to build an antiballistic missile system, and these hearings were prominently featured on national network news. At the height of the controversy, the Gallup Poll asked whether its respondents had heard or read about the ABM program (Gallup, 1972, p. 2190). Thirty-one percent were willing to admit that they had not heard of the program. In all probability many of those who said they had heard of the proposal were simply too embarrassed to admit they had not, for an additional 29 percent of those polled stated that they had no opinion on the issue. Questions on the information level of the ordinary voter raise some of the most intractable issues in democratic theory, particularly on the question of whether a government or its leaders are ultimately accountable for the success or failure of policy.

Sears (1969) provides a good discussion of the evidence concerning the political knowledge of the average citizen. Suffice it to say that most of our

measures of information tap such issues as whether or not a person knows his elected leaders or knows basic facts about the constitutional structure of the government. Measures of knowledge about political controversies or practical politics are quite rare. For example, numerous surveys chart the growth of disapproval of Richard Nixon as the Watergate crisis unfolded. However, no published survey through 1974 has actually measured what knowledge people gained of the materials on the Watergate tapes or the nature of the charges of the House Judiciary Committee's impeachment report. Without such knowledge it is not at all clear what kind of impact, if any, the events had upon people's perspectives on government. This is simply one illustration of a general point, that in the main we have only the most tentative information about people's information.

Finally, we would want to know about the *context* of a belief. Taken together, how do the opinions of individuals aggregate to form a *public* opinion? How far is a particular opinion from the modal view? What is the character of aggregated opinions as a distribution? And as Key (1964) and Dahl (1972) asked, What distribution denotes a consensus? polarization? a constitutional crisis?

To inquire of the context of a belief is also to ask about the relationship between a person's beliefs and those of political leaders. Does a person see significant differences in policies between a Democratic candidate and a Republican? That, of course, depends on the attitudinal distance of the two candidates from the observer's own position. Or, as someone once put it, the difference between a Methodist and a Baptist may not seem like much to a Buddhist, but it's like night and day to the Baptist.

Another aspect of context includes beliefs about *other* people's beliefs. Allport (1924) and Katz and Schanck (1938) have underlined the importance of *pluralistic ignorance*, the false ideas we have of the attitudes of others. Allport warned against the "illusion of universality" which sustains people in action through the belief that they are supported by an army of like-minded supporters. Allport spoke of pluralistic ignorance as an important source of prejudice. It reminds us that in an "other-directed" culture, the perception of beliefs of others may have as great an influence on our behavior as our own beliefs. Madison, in *Federalist* No. 49, knew that people are hesitant to act on opinions unless they perceive that others support them. In the context of a warning against Jefferson's proposition in *Notes on the State of Virginia* that constitutions should be amended by conventions of the people, Madison remarked:

If it be true that all governments rest on opinion, it is no less true that the strength of opinion in each individual, and its practical influence on his conduct, depend much on the number which he supposes to have entertained the same opinion. The reason of man, like man himself is timid and

cautious, when left alone; and acquires firmness and confidence, in proportion to the number with which it is associated. (Cooke, 1961, p. 340)

Another component of the context of a belief has to do with the expectations in which the beliefs are grounded. We have expectations, for example, based on our perceptions of the fate of others in situations similar to our own. Theorists of *relative deprivation* tell us that satisfaction with one's material life is less determined by one's objective situation than it is by our perception of the situation of others. As an illustration, Stouffer *et al.* (1949, vol. 1, pp. 250-253) found that soldiers in units with few promotions were more satisfied with promotion opportunities than soldiers in units with a high record of promotions.

Our beliefs are also embedded in the context of what Lewin (1943) has termed a "time perspective." We see ourselves and the world in which we live against the backdrop of the past and the unfolding of the future. Thus, a person's policy beliefs are not simply a function of his preferences but also of the difference between what he prefers and what he foresees realistically as the probable course of his life. To take a hypothetical example, two white citizens of South Africa may share the same attitudes toward apartheid laws, but if one believes in the inevitability of African enfranchisement while the other does not, the two are likely to have very different stances on the liberalization of race legislation.

Political judgments. Quite often a survey solicits not so much a respondent's attitude as it does a judgment, for example, not what candidate does a person *want* to win in an election but who does he *expect* to win. Respondents will surely blur the distinction between an attitude and an appraisal, no matter how much the item wording may entreat him to suppress his own opinions. Nonetheless, judgments or appraisals are a distinct analytical category, and we should so treat them.

A clear example of a judgment measure is the "self-anchoring striving scale," developed by Hadley Cantril and Lloyd A. Free (Cantril, 1965 and Free and Cantril, 1968) for a series of studies in 18 countries between 1958 and 1964. Each person was asked to evaluate the nation's achievements in relation to his own hopes and fears for it. Thus, the scale is distinctive not only as a measure of judgment, but also because the anchor points of the index are provided by the respondent himself. Each person is first asked to describe "the best possible situation for our country," then "the worst." He is then shown a ten-point ladder with the explanation that the tenth or top rung represents the best he could hope for the nation and the bottom rung the worst. Having set the anchor points in terms of his own best hopes and worst fears, he then indicates where the nation stands on the ladder at present, where he believes it stood five years ago, and where he expects it to stand five years in the future. (It is worth noting that this is a time perspective of five years imposed on the

respondents. A different time frame might have produced quite divergent findings.)

Most people in most countries render favorable judgments on their country's progress. In the 18 countries surveyed by Cantril and Free, only once (in the Philippines in 1959) did a nation judge its present to be worse than its past. Notably, a recent survey by Albert Cantril and Charles Roll (1971) established the United States in 1971 as a second example of a country whose population has judged its present as inferior to its past. Such has been the growth of disaffection in the United States since the mid 1960s (Boyd, 1974, p. 181).

An example of a judgment measure that is part appraisal, part attitude and affect is the "presidential popularity" question, which Gallup has asked regularly since 1945: "Do you approve of the way (the incumbent) is handling his job as president?" Mueller (1973) has given us an interesting analysis of why presidents have always declined in the nation's estimate during office. He finds that the positive judgments suffer in times of increased unemployment (though good times do not necessarily benefit a performance rating). Flash crises produce a rallying to the president, but this kind of approval quickly fades. And presidents can always expect to suffer a loss of approval the longer they have served in office, because people seem to remember the failures of policy more acutely than the successes. Counterintuitively, Mueller found that when the several causes of unpopularity are considered, "the Korean War had a large, significant independent negative impact on President Truman's popularity of some 18 percentage points, but the Vietnam War had no independent impact on President Johnson's popularity at all" (Mueller, 1973, p. 277). This is but one of a number of arresting findings about an item that has been asked perhaps as many times as any other political question in any survey series, but which had never been subjected to rigorous analysis prior to that of Mueller.

Behavior. Many of us, if asked to state the primary uses of survey research, would respond quickly, "Why to measure opinions, of course." So prevalent is this assumption that we often use public opinion surveys and survey research as synonymous terms. So we pause here to emphasize that measures of behavior are invariably important components of surveys and to dispute in the process the common charge that surveys of behavior are necessarily unreliable.

Voting choice is a type of behavior we often survey. Butler and Stokes, in their analysis of *Political Change in Britain* (1969), offer us a sophisticated interpretation of the shifts in party support during the period from 1959 to 1966 that reveals the complexity of seemingly simple changes in voting behavior. In 1959 the Conservative Party strengthened the margin it had enjoyed over the Labour Party since 1951, emerging with a 107-seat advantage. In 1964 Labour won a narrow victory of 13 seats over the Conservatives. With a margin so slim and prospects for a bigger victory so bright, Labour called another election in 1966 and triumphed with a 109-seat victory over the Conservatives.

Butler and Stokes were prescient enough to have captured this period of change with a sophisticated, three-wave panel survey of the United Kingdom (excluding Northern Ireland): a first in the summer of 1963, a second after the autumn election of 1964, and a third after the spring election of 1966. Having asked the 1963 sample about its 1959 voting choice, Butler and Stokes had a single sample of voters who had been interviewed at three different points of a formative political period. By cross-classifying voting choices in one election with those of another, they could precisely estimate the several sources of change (Butler and Stokes, 1969, pp. 275ff):

1. *Replacement of the electorate.* The coming of age of a new voter cohort and the death of a political cohort since the preceding election.
2. *Differential turnout.* The changes caused by people who vote in one election but not the other.
3. *Circulation of the Liberals.* Those "floating voters" who supported the Liberal Party in one election and the Conservative or the Labour Party in another. (The Liberal Party almost doubled its popular vote from 1959 to 1964.)
4. *Circulation of minor-party supporters.* The vote for Irish, Welsh, and Scottish Nationalist, or Communist candidates.
5. *Straight conversion.* Those who voted for the Conservatives in one election and the Labour Party in another.

The analysis dispelled the notion that most electoral change results from the final category—the change in vote from one major party to another. From 1959 to 1964 the greatest change was based in the replacement of the electorate—the substitution by those who were voting their first time of those who had voted their last. In second place as a cause of change was the circulation of Liberal voters. Straight conversion ranked only fourth of five in its impact. In short, political change was as much actuarial as electoral, an engaging finding that would have been lost were it not for the ability of surveys to capture the character of behavior.

Incidentally, this analysis of the sources of the Labour Party's surge required that the sample estimates be as accurate as possible. In this particular case, the marginals of the change tables could be adjusted to reflect the actual percentage of votes received by each party according to official returns. Mosteller (1968) had previously suggested a method of adjusting the cell frequencies in a table in a way that preserves the existing statistical relationship within the table, but alters the cell and marginal frequencies to reflect a distribution known to be more accurate than the sample results. Axelrod (1972, 1974) used the same adjustment procedures in his calculations of the coalitions that comprised the Republican and Democratic parties from 1952 to 1972.

A second example of using surveys to measure behavior attests to the versatility of the technique. In the summer of 1966 the National Opinion Research Center surveyed for The President's Commission on Law Enforcement and Administration of Justice a full probability sample of 10,000 households in all parts of the United States. Criminal victimization in the United States was the subject of the survey, specifically a comparison of the incidence of crime as reported in the FBI's Uniform Crime Reports with crimes mentioned by the sample respondents. Ennis (1967), the study director, found that the official figures approximated the survey results quite closely for such crimes as homicide and vehicle theft. For other types of crime, the survey estimates greatly exceeded the official figures, robbery being twice as frequent as recorded and rape being four times as frequent. The figures are, we might say, arresting. The survey also included a thorough analysis of why victims did not report crimes against them. And the careful analysis of victims' common dissatisfaction with police treatment in cases that *were* reported suggests some of the reasons for the decisions of many others to keep the crimes secret. In 1973 the Law Enforcement Assistance Administration of the United States Department of Justice initiated what is projected to become a continuing National Crime Panel. Conducted by the Bureau of the Census, the survey includes a sample of 60,000 households and 15,000 businesses throughout the 50 states and the District of Columbia (United States Department of Justice, 1974). Together, these criminal victimization studies cast doubt on the usual premise that survey results provide "softer," less reliable data than other methods of data collection.

Finally, we could not close our discussion of surveys of behavior without reference to perhaps the most extraordinary survey ever undertaken: the Multinational Comparative Time-Budget Research Project. This study, directed by Alexander Szalai in collaboration with Philip E. Converse, Pierre Feldheim, Erwin K. Scheuch, and Philip J. Stone, is based on detailed time-budgets recorded by nearly 30,000 respondents drawn from urban and suburban samples in 1965 and 1966 in 12 countries in Eastern and West Europe, South America, and the United States. From detailed reports of the activities of a 24-hour day, the investigators compare the use that citizens of different countries make of time. The results are illuminating. There could be no more precise support for Dahl's argument (1961) that man is not a political animal than to note that the average American spends fewer than 12 seconds a day on political activities (Szalai, 1972, p. 557).

However, if man is not Dahl's *Homo politicus*, neither does he appear to be *Homo copulatus* or *Homo evacuatus*. Though activities requiring an average of fewer than three seconds were not recorded in the tables, the tables suggest that no one devoted a single moment to sexual intercourse or elimination—not even during the weekend! But perhaps this statistic can only suggest that surveys of behavior are not completely free from measurement error. As the authors so discreetly suggested in a footnote, "For obvious reasons, minor

unidentified gaps in the flow of daily activities were by fiat assigned to the 'personal care' category" (Szalai, 1972, p. 129).

Ideologies and belief systems. In the last decade surveys of ideologies and belief systems have been undertaken in some form in most of the industrialized countries of the world. Many of these studies have focused on the problem raised by Converse in his classic article "The Nature of Belief Systems in Mass Publics" (1964). Converse defines a belief system in terms of what he calls a "constraint or functional interdependence" among a person's attitudes. Converse distinguishes two different types of constraint, static and dynamic. By static constraint he means "the success we would have in predicting, given initial knowledge that an individual holds a specified attitude, that he holds certain further ideas and attitudes" (Converse, 1964, p. 207). In contrast, dynamic constraint "refers to the probability that a change in the perceived status (truth, desirability, and so forth) of one idea-element would *psychologically* require, from the point of view of the actor, some compensating change(s) in the status of idea-elements elsewhere in the configuration" (1964, p. 208). Summarizing Converse's long and sophisticated essay invites oversimplification, but in general he argues that a continental shelf exists between elites and ordinary voters in terms of having either constrained belief systems or even stable positions on discrete issues over time.

Converse's work has stimulated numerous critiques, which he evaluates in Volume 4 of this *Handbook of Political Science* (1975). Since our interest is only in arguing that belief systems can be usefully studied with survey research, we will not review the controversy in depth, except to note that Cobb (1973) provides a helpful review of this literature.

Cobb suggests that research has focused on six features of belief systems: (1) the degree of differentiation of beliefs, whether they are cognitively simplistic and undifferentiated or sophisticated and complex; (2) the intensity and commitment of ordinary people to politics; (3) the stability of beliefs, or the extent to which peoples' attitudes remain fixed over time; (4) the internal consistency of attitudes within a belief system, that is, the degree to which one can predict that people who are liberal on one attitude will be liberal on another; (5) the instrumentality of attitudes, or the degree to which beliefs predict behavior; and (6) the insulation of belief systems, or the degree to which ideas in an ideology are vulnerable to change from external pressures.

Arguments on all sides of these issues can be buttressed from evidence gathered through sophisticated survey designs, which suggests two conclusions regarding the usefulness of surveys in studying ideologies. First, the degree to which people possess constrained belief systems in both a static and a dynamic sense can vary substantially in the same political system over a short period of time. For example, Converse's evidence suggesting that voters are ideologically unsophisticated is based on election surveys of 1956, 1958, and 1960. However,

as so many have remarked so often, this was indeed a quiescent period in American politics, when politics seemed relatively unimportant to most people. The 1960s, however, brought forth a set of issues that engaged voters' attention, most notably the civil rights movement and the antiwar movement, and candidates such as Barry Goldwater, Lyndon Johnson, George Wallace, and George McGovern took distinct and widely divergent stands on these issues. Voters paid attention to these stands, evidently took cues from them, and responded by holding views on political issues that were both ideologically sophisticated and internally consistent (e.g., Field and Anderson, 1969; Boyd, 1972; Luttbeg, 1968; Nie, 1974; Pierce, 1970; Pomper, 1972; and RePass, 1971, among many others). Taken together these studies clearly document a remarkable change over a short period of time in the degree to which voters think about politics in ideological terms, a change in large part determined by the salience and character of political issues at the time and the clarity of the positions that leaders are taking on those issues. It is not the case that voters do not respond to political issues because they lack the basic cognitive skills to do so. Thus, what surveys disclose about the ideological character of voters' attitudes may say as much about the character of political controversies or the conduct of leaders as it does about immutable features of people's cognitive capacities.

None of this argument is intended to deny the undeniable—that highly educated people, people of high cognitive ability, are more likely to have structured opinions than are people of low cognitive development. For example, Williams and Wright, using a factor analysis conducted separately for each educational stratum, found that “the organization of opinions appear[s] more strongly delineated in the college educated segment” (1955, p. 563). Similarly, Campbell et al. (1960) showed that ideological “levels of conceptualization” were much more common in the college stratum than among lower education groups. This suggests that the secular change toward increased education over the last 20 years might be expected to have a significant effect on the degree to which ordinary citizens possess structured opinions about politics. We commonly think of education as an experience that induces us to think in more sophisticated and complex ways about politics. Yet, we must note that Nie (1974) found that the growth of more structured opinions from 1956 to 1972 was as great in the stratum that had not attended college as in the stratum that had. Thus, it seems clear that it is both the changeable political climate that people confront, as well as their more enduring cognitive abilities, that determine the degree to which they possess structured opinions about the political world.

Second, surveys may be a poor device for exploring what we ordinarily mean when we use the term ideology. By ideology we usually have in mind those abstract social and political values that people can use to justify a set of attitudes or policy preferences. The standard interview schedule, even with its complement of open-ended questions, is not well suited to the in-depth probing

exploited by Lane (1962) and Smith, Bruner, and White (1956) to elicit the terms in which a person justifies his beliefs. With the exception of the literature on the "levels of conceptualization scale" introduced in Campbell's *The American Voter* (1960), most studies of belief systems have focused on responses to closed questions and have analyzed the degree of internal consistency in a person's attitudes, asking, for example, whether one can predict that a person who is liberal on one issue will be liberal on another. Where consistency exists we have tended to assume that it is because people's attitudes are integrated into a more encompassing belief system. Where it does not, we have inferred that no such ideology exists. But we have not directly measured the ideological justification; the basis for our interpretation is little more than a statistic, the degree of intercorrelation of beliefs across a sample of respondents. By judging attitudinal consistency in this manner, we require not only that a person have an ideological justification for his beliefs but that the respondents in the sample share the *same* justification.

Axelrod (1967) neatly demonstrated an alternative approach. In his analysis of a 1956 American survey, he found a group of citizens (mainly the politically active and well-educated) whose policy views were consistently liberal or consistently conservative. But he also found another group, commonly less-educated nonvoters, who took consistent positions in support of social welfare policies, of firing suspected communists from public service, and of an isolationist foreign policy. This latter set of beliefs he labeled "populism" because its advocacy of a strong governmental role in domestic affairs, combined with isolationist and anticivil libertarian beliefs, corresponded to Hofstadter's (1955) interpretation of the attitudes of the American Populist movement of the 1890s.

The moral is plain. Axelrod found evidence of consistent beliefs where others did not because he was willing to drop the assumption that all respondents in the sample must evaluate policies in terms of the same dimensions. Instead, he looked for consistency among subgroups of citizens whose social and political character was more homogeneous and who might be expected to share similar perspectives on politics.

As Axelrod's example suggests, the character of belief systems is grounded in the pattern of a nation's social cleavages. In some countries, notably the Netherlands, Belgium, India, and South Africa, the dominant cleavages of class, region, and religion all coincide with ethnic lines. Thus in such countries, there tend to develop separate trade unions, voluntary associations, businesses, newspapers, and religious organizations within each ethnic group. The Dutch word *verzuiiling* has become a standard term for describing these systems in which the major social cleavages parallel or reinforce one another, rather than crosscut each other (Lipset and Rokkan, 1967, p. 15; and Lijphart, 1968). When people are embedded in social groups in which everyone tends to have the same religious, ethnic, organizational, and perhaps even class loyalties, it would not be surprising to find substantial attitudinal consistency across a series of political

issues, for each of his reference groups reinforces similar political beliefs. In a country of crosscutting cleavages, however, people are constantly placed in positions of cross-pressures because their social groups tend to be more socially and politically heterogeneous. For this reason several of the recent surveys of political beliefs in various countries have focused on the linkages between patterns of social cleavages and the consistency of political attitudes (Barnes, 1971; Barnes and Pierce, 1971; Converse and Pierce, 1970; and the various studies in Lipset and Rokkan, 1967).

Studies of national character. To this point we have been discussing relatively limited domains—affect, attitudes, action. In contrast, surveys are often interested in the most comprehensive analyses of a nation's modal character structure. When such studies are limited to one society, they run the risk of drawing inferences about the distinctiveness of national character on the basis of "pseudo-cross national designs" (Hyman, 1964). That is, the analyst compares the empirical results of a survey in one country to the results of an imagined survey in a second. To avoid this problem there was a widespread, one could almost say worldwide, interest in the 1950s and 1960s in truly cross-cultural surveys of personality in which uniformities and differences in character types could be empirically demonstrated.

One approaches the topic of national character studies as one does a wasp nest, knowing that no matter how careful a person is, he will inevitably be stung. Fortunately, there exists a protective netting—in the form of numerous works aimed precisely at clarifying the difficult problems that cross-cultural research creates, and we will take refuge by citing those works that are most helpful in coping with them. At the beginning of the list is the annotated bibliography compiled by Frey and his associates Stephenson and Smith (1969), who have abstracted over 1600 articles on comparative surveys published up to 1967, classified by topic and area of the world.

Several works delineate the alternative research designs for cross-cultural surveys, including Frey's "Cross-Cultural Survey Research in Political Science" (1970), Przeworski and Teune's *The Logic of Comparative Social Inquiry* (1970), and Hyman's "Strategies in Comparative Survey Research" (1975). Needless to say, our definition of surveys in terms of "explicit, standardized procedures yielding quantitative measurements" confronts nearly intractable problems when interview schedules must be translated into other languages and when research procedures must be equivalent across cultural boundaries. It would be fatuous to seek solutions to linguistic and cultural differences with literal translations of instruments and nominal identity of research procedures. Rather, the cross-cultural researcher strives for functional or conceptual *equivalence* of procedures. Some of the more helpful sources, in addition to those named above, are Hymes, "Linguistic Aspects of Comparative Political Research" (1970); Przeworski and Teune, "Equivalence in Cross-National Research"

(1966–1967); and Strauss, “Phenomenal Identity and Conceptual Equivalence of Measurement in Cross-National Comparative Research” (1969). Specialized problems of equivalence are discussed by Haller and Lewis (1966), Gough and Di Palma (1965), Tiryakian (1958), and Schuman (1966).

Studies of *national character* are a special type of cross-cultural investigation, distinctive in their search for modal personality types arising in particular sociocultural systems. In *The Handbook of Social Psychology*, Inkeles and Levinson carefully state the issues that studies of national character pose:

The concept of national character is an important but problematic one in the social sciences. It has been strongly rejected in the hereditarian or racist forms in which it was couched by earlier writers. Seen in more modern perspective, however, it poses fundamental problems for social-scientific theory and research: To what extent do the patterned conditions of life in a particular society give rise to certain distinctive patterns in the personalities of its members? To what extent, that is, does the sociocultural system produce its distinctive forms of “social character,” “basic personality structure,” or “modal personality”? Further, what are the consequences, if any, of this patterning in personality for stability or change in the societal order? (Inkeles and Levinson, 1969, vol. 4, p. 418)

Together with DeVos and Hippler (1969), the essay by Inkeles and Levinson is a valuable guide to this literature.

Controversies in Attitude Research

In our discussion of attitudes and behavior, we deferred examination of issues that continue to be the subject of controversy. Prominent among these are (1) the reliability of attitude measures in light of the sensitivity of responses to slight changes in item wording, (2) the degree to which people’s behavior conforms to attitudes expressed in interviews, and (3) the degree to which responses to many items can actually be said to measure political attitudes, when respondents tend to answer such questions in almost random patterns over time (the so-called “non-attitude” argument). Each of these issues poses difficult problems for the survey analyst, both in the design of measures and the interpretation of results.

Response sensitivity to item wording. Perhaps the first warning the survey initiate receives is that subtle changes in the wording of attitude items often alter the frequency of a response by as much as 20 percentage points. Experienced analysts wince at reading a statement that, for example, 80 percent of a sample supports gun control legislation or that 55 percent endorse legalized abortion. They know that a differently worded item could easily turn these majorities into minorities. Indeed, experiments on the effects of item wording were a core concern of those who were the pioneers of modern survey techniques

in the 1930s and 1940s (Gallup, 1944; Cantril, 1947; Stouffer et al., 1949; and Payne, 1951).

Philip Converse and Howard Schuman (1970) offer us a recent example of this old problem. In June 1969 the Gallup Poll asked a national sample the following question on Vietnam: "President Nixon has ordered the withdrawal of 25,000 United States troops from Vietnam in the next three months. How do you feel about this—do you think troops should be withdrawn at a faster rate or slower rate?" (The response "same as now" was not presented, but it was accepted if volunteered.) In September–October 1969, the Harris Poll asked a similarly worded item: "In general, do you feel the pace at which the President is withdrawing troops is too fast, too slow, or about right?" The responses to the two items were the following:

<i>The Gallup Poll</i>		<i>The Harris Poll</i>	
Faster	42%	Too slow	28%
Same as now	29	About right	49
Slower	16	Too fast	6
No opinion	13	No opinion	18
	<hr/> 100%		<hr/> 101%

The differences in the poll results are striking. One could interpret the Gallup Poll as a plurality in favor of faster withdrawal. Yet the Harris Poll, with a very similar item, shows a plurality in support of administration policy. Why is there a 20 percentage point difference in the two items? Obviously the answer is that the Gallup Poll did not suggest to its sample the response alternative, "same as now." Yet one would hardly expect this change to affect people's stated opinions on an issue so salient as the war, measured at a peak point of its intensity.

Confronted with item wording effects such as these, the prudent analyst is usually given the following advice (Stouffer et al., 1949, vol. 1, p. 45; and James Davis, 1971): Avoid imputing any absolute validity to the total distribution of responses because wording changes can shift the modal response on an attitude dimension. However, one can usually assume, one is told, that the *relative ordering* of people in each response category is still reliable, e.g., that most of the 28 percent in the Harris Poll who supported a faster withdrawal of troops would have also answered the Gallup item by advocating a faster withdrawal. Thus, the two items still measure the same content domain. They simply slice that opinion distribution into groups of somewhat different sizes.

If it were true that the two items measure the same underlying attitude, then it should be the case that the items will have a similar statistical relationship to a third variable such as education. Thus, for either sample one could

formulate a generalization such as "the more educated a respondent in each sample is, the more likely it is that he favors a faster withdrawal of troops." Such a generalization would require that respondents of the same educational level interpret the meaning of the differently worded items in a similar way. However, not even this assumption is always valid, as Schuman and Duncan (1974) reveal in a subsequent reanalysis of these same two Vietnam items. Table 2 displays the quite different statistical relationships that responses bear to education.

TABLE 2 Percentage distribution on withdrawal from Vietnam questions by education: Gallup and Harris Polls*

	Advo- cates faster with- drawal	Advo- cates slower with- drawal	Advo- cates present rate†	No opinion	Total %	N
Gallup						
College (13 and over)	42	16	31	11	100	300
High school (9-12)	42	16	30	12	100	765
Grade school (0-8)	41	19	24	16	100	435
						1500
Harris						
College (13 and over)	36	4	42	18	100	320
High School (9-12)	27	7	51	15	100	816
Grade school (0-8)	18	5	61	16	100	468
						1604

Source: Schuman and Duncan, 1974, p. 239.

* See text for wording of Gallup and Harris questions. Responses to the two questions have been equated and relabeled here in terms of intended meaning; for example, "faster" for Gallup question and "too slow" for Harris question are both considered "advocates faster withdrawal."

† This alternative is explicitly presented to respondents in this Harris question but not in the Gallup question.

In the Gallup Poll, those with a college education are more likely than those with a grade school education to favor the existing rate of troop withdrawal. In the Harris Poll this positive relationship becomes negative, as those with a grade school education give more support to the administration policy than the college-educated do. As Schuman and Duncan conclude, "the claim is often made that orderings and associations are invariant in the face of changes in question wording. There are in fact few demonstrations that this assumption is generally true" (1974, p. 236).

Indeed, how can the assumption ever be justified? The 1968 Survey Research Center election study offers us one example that lends it credibility: the effect of item wording on the venerable political efficacy index. The index, which we introduced previously in our discussion of affect, was formulated in the Survey Research Center election survey of 1952, and the questions were phrased in a format common to that period. That is, the items were declaratory statements, with the respondent asked to state his agreement or disagreement with the assertion. Table 3 lists these items in their original versions.

Items stated in this "agree-disagree" format are often invalid measures because of the well-known tendency of many people, especially less-educated respondents, to agree to any statement regardless of its content. This particular form of bias is known as *acquiescence response set*. These items are twice biased because every item is worded such that an agree response denotes a feeling of political powerlessness. Thus, respondents who score high on this powerlessness index are made up of some unknown quantity of people who actually feel powerless, plus another unknown number who are obligingly agreeing with the statements.

To avoid such bias the Michigan Survey Research Center began to drop the agree-disagree format in favor of item wordings in which the respondent is asked to choose among alternatives stated in the question, rather than simply agree or disagree with a statement. The revised version as worded in Table 3 was therefore included on the 1968 preelection survey. Yet, because of the interest of many scholars in keeping the original items for time series analysis, the original versions were included in their usual place in the postelection survey. Thus, the two versions of the index were asked of the same respondents but in interviews separated in time by one to five months. Given the fact that the sample is the same, any substantial variation in responses is likely due to the effects of item wording. The two versions present, then, an attractive experiment in item-wording effects.

As Table 3 clearly indicates, the impact of item wording is considerable. Indeed, the majority position becomes the minority position on three of the four items. Were it not for the presence of the original set we might have inferred extraordinary changes in the public feelings of political efficacy, when in fact the source of much of the change would have been in the typewriters of the items' authors. Item 4 provides the most extreme variation; the proportion expressing feelings of powerlessness is almost 35 percentage points higher than in the original version.

Interestingly, this case supports the maxim that differently worded items may still measure the same content domain even though the items may elicit very different frequency responses. Despite wide variations in the pattern of responses to each pair of items, the responses themselves are highly intercorrelated. That is, those who expressed feelings of powerlessness in one version of the test had a high probability of expressing such feelings in the other. The sta-

TABLE 3 Frequency of responses to two versions of the SRC political efficacy items, 1968.

Original version	Revised version
1. "I don't think public officials care much what people like me think." (N = 1312)	1. "Would you say that most public officials care quite a lot about what people like you think, or that they don't care much at all?" (N = 1475)
Agree 44%	Don't care 40%
Disagree 56%	Care 60%
.61*	
2. "Voting is the only way that people like me can have any say about how the government runs things." (N = 1323)	2. "Would you say that voting is the only way that people like you can have any say about the way government runs things, or that there are lots of ways that you can have a say?" (N = 1493)
Agree 43%	Voting is the only way 58%
Disagree 57%	Lots of ways 42%
.69*	
3. "Sometimes politics and government seem so complicated that a person like me can't really understand what's going on." (N = 1336)	3. "Would you say that politics and government are so complicated that people like you can't really understand what's going on, or that you can understand what's going on pretty well?" (N = 1521)
Agree 71%	Can't understand 44%
Disagree 29%	Can understand 56%
.68*	
4. "People like me don't have any say about what the government does." (N = 1331)	4. "Would you say that people like you have quite a lot of say about what the government does, or that you don't have much say at all?" (N = 1507)
Agree 41%	Don't have much to say 75%
Disagree 59%	Have a lot to say 25%
.67*	

* Gamma coefficient measuring association between two versions of items.

tistical relationships between the two versions of the test are invariant over changes in the item wordings, for the correlations are high and steady within a fairly narrow range from .61 to .69.

Moreover, as Table 4 indicates, responses to the two versions of the index are quite similarly related to the respondent's level of education. In both versions feelings of political powerlessness are concentrated among people with less education. We can also see the effects of acquiescence response set on the original version. Because the less educated also tend toward acquiescence, the disposition of the less educated to feel powerless is reinforced by the response set. Thus, in each of the four items the association of the original items with education is higher than it is with the revised version.

These two examples, attitudes on the Vietnam War and feelings of political powerlessness, offer us illustrations of the range of item-wording effects we

TABLE 4 Relationship of two versions of political efficacy items to Education, 1968: proportion giving inefficacious responses.

Item	Education of Respondent					Gamma
	Grade school or less	Less than high school graduate	High school graduate	Some college	College graduate	
1. Public officials don't care						
Original version	76%*	59	35	28	20	.51
Revised version	62%	44	38	35	27	.25
2. Voting is the only way						
Original version	87%	73	53	44	23	.54
Revised version	83%	68	58	45	27	.46
3. Politics is too complicated						
Original version	88%	86	69	57	46	.50
Revised version	70%	55	41	31	20	.43
4. People don't have say						
Original version	71%	52	36	28	21	.43
Revised version	86%	81	74	69	61	.29

* Each cell entry is the proportion of respondents expressing a feeling of political powerlessness. Because the response categories are dichotomous, the proportion expressing feelings of political efficacy is the difference between the figure and 100 percent. The items are those of Table 3.

should anticipate. In the case of war attitudes, the form of the question affected the frequency of responses to the items and modified the statistical relationships between the war attitudes of the respondents and their education. In the case of political efficacy, item wording had substantial effects on the frequency of responses but minimal impact on their statistical relationships to another variable, education. Converse and Schuman (1970) also provide other examples in which different wordings of items produced no significant effects of either type. Thus, there are no uniform rules that the researcher can confidently apply to the problem. Perhaps the best advice is given by Stouffer *et al.* (1955) long ago, and elaborated by Donald Campbell and Fiske (1959) and Schuman and Duncan (1974): whenever possible measure a concept with as many different items as practicable and analyze separately the relationships of these multiple indicators to other variables. Similarly, Noelle-Neumann (1970) argues for the routine use of split ballots to control for question wording effects. Though multiple indicators consume scarce questionnaire space and split ballots complicate data analysis, each technique provides a basis for confidence that the relationships one observes are not simply a result of an arbitrary wording of an attitude item.

Even when we find that pleasing situation in which changes in item wordings do not disturb statistical relationships with other variables, the effect of item-wording changes on response frequencies remains troublesome. This is particularly true when we wish to examine changes in attitudes over time. Duncan (1969, p. 28) concisely states the dilemma: "When a study has become old enough to be interesting as a base-line for change measurement, it is likely also to be old enough to have used techniques considered outmoded in some respects. Hence there is a great temptation to substitute improved measurements instead of achieving serious replication [of original measures]." Duncan suggests that researchers should generally use improved measures even for time comparisons, but that whenever possible they should include the old measures on questionnaires along with the new. By this method, the new measures can be calibrated or spliced together with the old one, preserving our investment in time series data.

The United States Bureau of Labor Statistics' measure of unemployment illustrates nicely Duncan's advice. (To be sure, measurement of the labor force is not strictly a "political" variable, but, of course, it has profound political impact.) For many years it was known that the measure of unemployment was defective. However, to have changed it might render impossible the kinds of longitudinal analyses for which the time series on unemployment is so important. However, applying new and old measures to the same years allowed the two indexes to be calibrated in such a way that the unemployment rate for years past could be estimated as though it had always been measured using the newly revised series (Ducoff and Hagood, 1947, pp. 5-7, 28-29, 89).

Unfortunately, the calibration of measures does not always avoid the con-

flict between the need to correct deficient measures and the desire to protect a time series for longitudinal analysis. This dilemma is apparent in the changes that the Survey Research Center has made in many of its attitude items.

From 1952 through 1960 many Survey Research Center items other than political efficacy were phrased as statements to which the respondents were asked to agree strongly, agree, disagree, or disagree strongly. As we discussed in the example of the political efficacy items, this format is subject to various biases, including acquiescence response set. Presumably, because the 1956, 1958, and 1960 studies were a panel design with a primary emphasis on longitudinal analysis, these measures were retained through 1960 in spite of their probable biases. With the return to a standard cross-section sample in 1964, many of the old items were changed to the format illustrated by the revised political efficacy items of Table 2, in which respondents must choose between response alternatives embedded in the items. In most cases the old item formats were not included for the purposes of calibration, at least not in the final interview schedule. Thus, researchers who wish to use the now substantial series of Survey Research Center election studies for longitudinal analyses should assure themselves that the changes they observe are not artifacts of altered item wordings.

Nie's (1974) study of belief systems provides an illustration of the pains a careful researcher should take to eliminate the possibility that item-wording effects contaminate measures of attitude change. By comparing responses to the Survey Research Center items with differently worded items from a National Opinion Research Center (NORC) sample, Nie found that the revision of the Likert format for measuring opinions on public policy issues did not prevent their use in time series analysis. In other cases (the political efficacy items in Table 2 are a good example) item-wording effects are of such a magnitude that time series analyses combining the original and the revised versions would be difficult. Even though the relationships of the efficacy items to variables such as education were not disturbed, the wording effects were so large that the new version could not be spliced with the original. Evidently, the Survey Research Center decided that scholarly interest in maintaining the time series on the efficacy measures outweighed the advantages of substituting refined measures; in the 1972 survey the original items were included and the revised items discarded.

The nexus of attitudes and behavior. A second controversy concerns what Cobb (1973) calls the instrumentality of attitudes, that is, the degree to which people's behavior conforms to their beliefs. Although Deutscher (1966) argues that the issue remains important, his formulation of the controversy now seems somewhat quaint. Few would dispute that attitudes are sometimes poor predictors of actions. As Liska (1974) points out, the question has been recast in a more complex multivariate form: what conditions affect the correspondence

between beliefs and behavior? Reformulated, it simply becomes a theoretically interesting question to ponder why our attitudes and actions so often diverge.

1. A necessary first step in clarifying this controversy is to distinguish surveys of past and future behavior. If we seek to measure past behavior, then conceptually there is no uncertainty about whether the overt behavior on which the respondent is queried actually took place. Presuming the item is itself unambiguous, all that is at issue is the accuracy (memory error) and honesty (bias) of the response.

As Sudman and Bradburn (1974, p. 67) note, psychologists have rather accurately defined the distribution of memory errors. Sudman and Bradburn distinguish two types of memory errors in surveys. The first is *omission error*, in which a respondent simply forgets certain of his past actions. The second is *telescoping error*, when time is compressed and an event is remembered as occurring more recently than it actually did (for example, a respondent reporting a case of criminal victimization in the preceding month when in fact the crime occurred two months in the past).

Sudman and Bradburn specify the amount of response error contributed by omission and telescoping with some precision. Because the validity of survey reports can be measured in many cases, it is merely tendentious to argue that inaccuracy of responses generally invalidates surveys of past behavior. Rather what we wish to know is the magnitude of the errors and their distribution, random or systematic, across groups of respondents. We also wish to know how such errors can be reduced by interview procedures such as aided recall and open-ended questions. Sudman and Bradburn provide exactly these helpful types of guides.

Errors due to bias are often more intractable than random memory errors, since biased responses include systematic distortions into data analysis. Survey researchers have long been aware of the understandable disposition of people to report *socially desirable* responses, that is, what they wish the interviewer to believe about themselves. For example, during World War II the Division of Surveys of the Office of War Information surveyed respondents who were known to have redeemed war bonds in the preceding week. Seventeen percent of the sample denied having redeemed the bonds, an act considered at the time to be harmful to the war effort (Hyman, 1944-1945).

There has come to be a substantial literature devoted to estimating the magnitude of bias in surveys, for example, Dinerman (1948), Parry and Crossley (1950), Zitter and Starsinic (1966), Cahalan (1968-1969), and Clausen (1968-1969) on the systematic inflation of survey reports of voter turnout rates. Substantial efforts are being made to reduce such biases. One of the more exotic examples is the randomized response model, in which a respondent is presented with paired questions (Abernathy *et al.*, 1970; Folsom *et al.*, 1973). One is an

innocuous item for which the population parameter is well known; for example, "I was born in the month of April." The second item in the pair is a threatening one, likely to evoke biased responses; for example, "I was pregnant at some time during the past 12 months and had an abortion which ended the pregnancy." A random process dictates to the respondent which of the two questions she should answer. The interviewer, however, does not know whether a "yes" response is an answer to the first or the second question, thus assuring the respondent some degree of privacy of her views. No response can ever be attributed to particular individuals in the sample, but because the proportion of people born in particular months is known, researchers can accurately specify the incidence of actions that respondents would otherwise be hesitant to admit in a standard survey. In general, the more anonymous survey respondents perceive themselves to be, for example, through the use of telephone or self-administered questionnaires rather than face-to-face interviews, the lower will be the response bias for items containing a socially desirable answer (Sudman and Bradburn, 1974, p. 66).

Surveys probing future behavior present an entirely different conceptual problem from surveys of the past. It is possible to measure the direction and magnitude of response error for past behavior because one can construct validating tests for which the actual behavior of the respondent is known. In contrast, there is an irreducible element of uncertainty about reports of intention to take some action in the future, for there are always contingencies that may intervene between an intention to act and the action itself. Epistemologically, a respondent cannot lie to an interviewer about an act that has yet to take place. How indeed can we ever know how we will behave until actually confronted with a situation? All a survey can do is to try to make a necessarily hypothetical event realistic for the respondent. For example, in the midst of a September campaign Gallup does not ask his respondents how they will vote in the coming November election. Rather, he asks them how they would vote if the election were held on the date of the interview. Strictly speaking, then, most polls do not predict elections; they merely measure people's voting intentions at the time of the interview. In the 1948 United States election and the 1970 United Kingdom election, pollsters came to grief for their failure to measure late campaign surges and to predict accurately the turnout rate. But the polls were not necessarily inaccurate on the vote intentions measured during the period of the interviews (Mosteller *et al.*, 1949; and Abrams, 1970). In short, the use of surveys as a tool to forecast future events is often a misuse of survey methodology. The failure of such forecasts is less a criticism of surveys than it is a failure of judgment on the part of the interpreter of the survey. Dollard's (1948) codification of the conditions under which opinions predict behavior remains an important theoretical statement on the reasons why our intentions so often fail to predict our future actions.

We have reason for confidence in surveys of behavior in an intriguing finding of Sudman and Bradburn: the more salient the item to the respondent, the more likely he is to report his behavior accurately in a survey (1974, p. 39). This suggests that people's attitudes on controversial issues may often be fairly reliable guides to their subsequent actions. A convincing test of this hypothesis was executed as a component of the 1969 Detroit Area Study, directed by Irwin Katz and Howard Schuman and reported by Brannon *et al.* (1973). The first phase of this ingenious experiment consisted of a survey of people's attitudes on open-housing legislation. On the basis of the survey, the respondents were placed in two categories, one supporting a nondiscrimination open-housing proposal and a second group supporting an owner's right to sell his house to whomever he chooses.

The action phase of the experiment followed three months later, when a group of graduate students posing as members of an organization of concerned citizens returned to each of the original respondents seeking endorsement of a petition supporting *either* open-housing or owner's rights, corresponding to his previously expressed attitudes. Eighty-five percent of the owner's-rights advocates signed the owner's-rights petition. Seventy percent of the open-housing advocates signed the open-housing petition. In contrast, among a control group of owner's-rights advocates who were presented with an open-housing petition, 78 percent *refused* to sign. On a controversial issue such as race policy, people not only expressed their attitudes to a sympathetic interviewer, they were also willing to commit themselves by name to a position on a document that, so far as they knew, could have been broadly available to the public.

2. Another element affecting the correspondence of attitudes and actions is the degree to which the attitudinal measures capture the genuine complexity of social life. This means, as Rokeach (1968) and Liska (1974) argue, that we must measure specific and often contradictory beliefs about concrete situations as well as those more abstract principles surrounding some controversy. Schuman (1972) provides an example from the 1969 Detroit Area Study that illuminates this point. As Table 5 shows, 85 percent support the principle of nondiscrimination in hiring (item C). Yet, 39 percent agree to discrimination if it is said to be necessary for the harmony of the firm (item A), and 50 percent agree to it if it is the preference of a majority of whites (item B).

Schuman argues that the different responses to the three items reflect the very real situations of our lives. Many of us will support principles stated in the abstract. Real situations, however, usually pose conflicts of principles, just as items A and B pose conflicts between belief in fairness on the one hand and belief in interpersonal harmony and the dominance of majority preferences on the other. When items are written to simulate these clashes of principle, surveys can tap the genuine ambivalence that most people feel when principles they endorse are in conflict and can elicit responses that are likely to bear a close

TABLE 5 Racial attitudes on an abstract principle versus attitudes on a concrete situation

Suppose a good Negro engineer applied for a job as an engineering executive. The Personnel Director explained to him: "Personally I'd never given your race a thought, but the two men you would have to work with most closely—the plant manager and the chief engineer—both have strong feelings about Negroes. I *can* offer you a job as a regular engineer, but *not* at the executive level, because any serious friction at the top could ruin the organization."

A. Was it all right for the personnel director in this case to refuse to hire the Negro engineer as an executive in order to avoid friction with the other employees?

1. Yes	39%
2. No	56
9. Other, D.K., N.A.	5

100(640)

B. Should the personnel manager have asked the other men how they would feel about working with a Negro engineer and then made his decision on the basis of their wishes?

1. Yes	50%
2. No	45
9. Other, D.K., N.A.	5

100(640)

C. In general, do you think employers should hire men for top management without paying any attention to whether they are white or Negro?

1. Yes	85%
2. No	12
9. Other, D.K., N.A.	3

100(640)

Source: Schuman, 1972, p. 348.

The sample was an area probability cross section of white adult heads and wives of heads of household in the metropolitan Detroit area. The response rate was 78 percent, with a final *N* of 640 individuals.

correspondence to what our behavior would have been in a real situation. As Hyman once observed, "If our aim is to predict a given kind of behavior in a given social setting, we should design our tests so that they incorporate the fundamental aspects of the setting into the test" (1949, p. 40).

Schuman's example of clashes of principles in hiring practices clarifies a dispute over the interpretation of two influential political studies by Prothro and Grigg (1960) and by McClosky (1964). Both had discovered an apparent inconsistency in Americans' support of democratic principles. Although Americans tended to endorse democratic principles stated in abstract terms, they

qualified that support substantially when presented with a specific instance of the application of those principles.

For example, respondents in both surveys endorsed the principle of free speech, but large numbers would not give that right to communists, atheists, and other unpopular groups. In the context of the cold war atmosphere of the late 1950s, when both surveys were conducted, it is perhaps not remarkable that people might feel that free speech for communists posed a conflict of principles between support for free speech and a belief that some political groups might pose a threat to that same democratic principle. Similarly, people may have felt a conflict of principles on the question of permitting antireligious advocacy in community forums and schools. Americans, as many others, have never really believed that questions of fundamental values (such as religion) are best settled in the marketplace of ideas. Rather, schools and churches in communities have often been charged with a responsibility for inculcating an appreciation of (most often) Christian doctrine, just as instruction in civics has long been fused into secondary school curricula. Thus, it should not surprise us that we find genuine disagreement on the rights of religious dissenters. We will return to the interpretation of these studies in a later discussion on the influence of survey research on political theory. It is sufficient here to note that the Prothro and Grigg and the McClosky studies provide evidence that survey questions can incorporate realistic conflicts between competing principles.

3. Finally, Liska (1974) emphasizes the impact of social support on the correspondence of attitudes to actions. When social support is congruent with attitudes, the correspondence is of course accentuated. When group pressures run counter to attitudes, social pressure may outweigh beliefs in affecting behavior. Liska's discussion of this point centers principally on the somewhat esoteric matter of whether social support and attitudes combine additively or interactively to affect actions. However, we can extend his point by noting the extensive literature stressing the importance of our location in social groups for our beliefs and behavior.

Primary groups—"those small, face-to-face, solidary, informal and enduring coteries that we commonly experience as family, friendship and occupational peer groups"—possess an "extraordinary capacity for rewarding conformity and punishing deviation" (McClosky and Dahlgren, 1959, pp. 757, 759). In their survey of the literature as of 1959, McClosky and Dahlgren list some of the important ways in which primary groups influence behavior and beliefs:

Almost every major voting study furnishes additional proof that primary groups are essential links in the complex process by which political norms are indoctrinated and party preferences implanted. They find, for example, that members of the same primary groups characteristically vote alike, think alike on issues, and affiliate with the same party; that voters in doubt about whom to vote for usually resolve their indecision by embracing the

political preferences of their friends; that approximately three out of four young people vote as their parents do; and that the more uniform a group's political outlook the firmer the voting intentions of its members. Homogeneity of opinion among primary group members also affects voting turnout and the level of political curiosity. People who disagree with their families or friends about politics are less apt to vote and less likely to develop or to retain an interest in politics. But primary groups may help to reinforce habits of participation and interest as well as to inhibit them. Patterns of participation, as one study concluded, are "contagious"—likely to be active when voters belong to politically aware groups and apathetic when they belong to politically indifferent ones. (McClosky and Dahlgren, 1959, p. 758)

Many of these findings derive from the sociological focus of Columbia's Bureau of Applied Social Research, which generated *The People's Choice* (Lazarsfeld, Berelson, and Gaudet, 1944), *Voting* (Berelson, Lazarsfeld, and McPhee, 1954), and *Public Opinion and Congressional Elections* (McPhee and Glaser, 1962), and which stimulated replications in the United Kingdom, *Straight Fight* (Milne and Mackenzie, 1954), *How People Vote* (Benney, Gray, and Pear, 1956), and *Marginal Seat* (Milne and Mackenzie, 1958). Since 1959 the emergence of the field of political socialization has added to our knowledge of the impact of family and peer groups on behavior (e.g., Hyman, 1959; Greenstein, 1965; Easton and Dennis, 1969; Hess and Torney, 1967; Jennings and Niemi, 1974; and Sears in Vol. 2 of this *Handbook*).

The role of primary groups is so well known that additional examples are superfluous. Less often noted are the effects of *secondary groups* and the *community* on our behavior, or what Ennis (1962) has termed "the contextual dimension in voting." One of the most intriguing concepts from this tradition of research is the concept of the *breakage effect*, a term Berelson borrowed from horse racing circles. Berelson, Lazarsfeld, and McPhee observed (1954, p. 100) that people who are cross-pressured tend to resolve their conflict by voting with the community majority. For example, in Elmira, New York, if people whose friends and coworkers were 2:1 Republican, the vote divided three-fourths Republican. But if they were divided 2:1 Democratic, only half voted Democratic. The "breakage" for the Republicans reflected the party's dominance in Elmira at that time: "... the Republicans get more than their random share of the adjustment to a conflicting environment, because of the pervasive Republican atmosphere of Elmira that thus tends to perpetuate itself. The surrounding majority gets the benefit of the operation of cross-pressures" (Berelson, Lazarsfeld, and McPhee, 1954, p. 100).

Butler and Stokes used the breakage effect to explain why, in English elections, the greater the proportion of a social class in a constituency, the more politically homogeneous is that class (1969, p. 146). Tingsten had first observed

this phenomenon and labeled it "the law of the social centre of gravity" (1937, pp. 177-180, 230). Valen and Katz (1964, p. 149) and Katz and Eldersveld (1961, p. 13) confirmed what they termed this "clustering" effect. Apparently, when it is numerically dominant, a group has more success in reinforcing the political behavior of its members (cf. Putnam, 1966). In addition, Butler and Stokes also found that a clustering or breakage effect explained the anomaly of uniform national election swings in the United Kingdom (1969, pp. 303ff). One might presume, for example, that a national swing of one out of every five normally Conservative voters to Labour would result in very large percentage losses in heavy Conservative constituencies and small losses in dominant Labour constituencies where there would be few normally Conservative votes to gain. In fact, however, there tended to be uniform party swings across constituencies regardless of the previous balance of party supporters. Butler and Stokes explain the uniform losses by noting that the breakage effect tends to offset the extraordinary losses that a party might expect to incur in those constituencies where its dominance would otherwise provide a large pool of potential defectors.

In sum, the primary and secondary groups to which we belong have a major effect on our behavior. It should not surprise us, then, that our actions will oftentimes be more influenced by our perceptions of the attitudes of others than by our own. Indeed, this is what survey research has confirmed, and it cannot be taken as a criticism of surveys that our behavior so often contradicts our beliefs. Given that all of the above findings have been generated from survey evidence, it suggests, quite the contrary, that surveys can be relatively refined techniques for clarifying the complex interrelationships of attitudes and actions.

Do surveys at times measure "non-attitudes"? A final controversy in attitude research is raised by Converse's provocative study, "The Nature of Belief Systems in Mass Publics" (1964). In the course of an essay of considerable scope, Converse examined the question of the degree to which people tend to have stable opinions over time. On what he termed items of "pure affect" (such as party identification is for many people), he expected and found fairly high stability of opinions from 1956 to 1960. On questions that invoked attitude towards groups *qua* groups (e.g., race-related issues such as school desegregation), he found a moderate degree of stability. However, on some policy issues, especially those that were designed to measure political ideologies about the role of the government in social and economic affairs, there tended to be little consistency of responses over time. "Faced with the typical item of this kind, only about 13 out of twenty manage to locate themselves even on the same *side* of the controversy in successive interrogations, when ten out of twenty could have done so by chance alone" (Converse, 1964, p. 239).

Moreover, Converse found a special property in the pattern of attitude stability on this last class of policy issues. The correlation for responses from

1956 to 1958 and from 1958 to 1960 was about .3 in both instances. Thus, one might have expected that the correlation from 1956 to 1960 would not have exceeded .09. In fact, the correlation for 1956–1960 was also .3. One model that would explain such a pattern would be, Converse suggested, a “black–white” distinction in which there might exist “first, a ‘hard core’ of opinion on a given issue, which is well crystallized and perfectly stable over time. For the remainder of the population [in this hypothetical model], response sequences over time are statistically random” (Converse, 1964, p. 242).

Converse explored a number of attitudes for goodness of fit with this model and found only one—the role of private enterprise and government in housing and electrical power—that fit the model of random change at all well. However, Butler and Stokes (1969, p. 181) found the same pattern of response correlations on the question of British attitudes toward an independent nuclear force. One implication of such findings—that the responses of *some* people to *some* policy items are so random as to properly be called “non-attitudes”—is sufficiently controversial that reinvestigations were inevitable. Converse has explored the issue further in “Attitudes and Non-Attitudes: Continuation of a Dialogue” (1970) and in Volume 4 of this *Handbook*. Pierce and Rose (1974) have examined whether an alternate probability model might not also explain why there appeared to be random responses to some policy items over time.

We will not pass final judgments in the dispute between Converse and Pierce and Rose. The reader is encouraged to make his own. The controversy has served to highlight some cautionary maxims that no one, regardless of his position on the non-attitude controversy, would likely dispute.

First, to pose survey questions of people who have no interest in them is to encourage meaningless, random responses. Thus, Converse notes (1974, p. 651) that the 1956 survey “invited respondents who had no opinions on a particular issue to report that fact directly, instead of laboring to concoct some kind of meaningless data point for us.” He suggests that this use of filter questions to weed out random responses ran counter to the trend of that period in which interviewers were trained to probe for responses in order to minimize missing data. The use of such filter questions to ensure that people who do not have opinions on some issue are not asked for an opinion is old and honored advice (Gallup, 1947). Unfortunately, it is too often ignored.

Second, the debate highlights the political situations in which people’s opinions are likely to display the most extreme instability across time. Converse took pains to note that instability of responses can occur for two very different reasons: (1) measurement invalidity due to random responses, and (2) real attitude change. In the first case people’s responses are unstable because their answers are elicited on topics for which they have no genuine attitudes. In the second case real attitudes are in flux because the level of politics has become sufficiently heated that people reexamine their beliefs.

There is no easy way to distinguish between instability due to random factors and that due to genuine attitude change. Therefore, in order to test his "black-white" model, Converse had to choose from a limited set of items in a special period in history for which he could assume that observed response instability was not due to attitude change. The 1956-1960 period of American politics was one of those quiescent times in which there were few political events that excited people to reflect on their political beliefs. In addition, Converse had to select a *type* of issue on which people's attitudes, if they had them, were relatively settled. By choosing the ideological issue of the role of the state in housing and electrical power, Converse could assume that there was little real attitude change on such a question. In short, Converse never assumed that his "black-white" model with its somewhat elitist implications would apply to more than a restricted type of issue—and then only in particular historical periods. Therefore, Converse's example probably represents a limiting case for the existence of "non-attitudes."

By making explicit the exceptional conditions in which the "black-white" model applies, our attention is focused on the conditions that produce the maximal response instability due to genuine attitude change as opposed to random "non-attitudes." The period would be one in which politics reaches a peak of political intensity for ordinary citizens, and the items would tap issues that had only recently emerged as salient. The late 1960s and early 1970s are obviously such a period in American politics, particularly on the questions of civil disorder and the Vietnam war. We might well assume that anyone who maintained perfectly stable attitudes on these issues during this period was either the most rigid of true believers or oblivious to the world around him.

Bibliography on Measurement and Scaling

The literature on attitude measurement and scaling is both voluminous and arcane; an analysis of it would require another chapter of this length and would be redundant, given that so many specialists have devoted themselves to the task (for example, Scott, 1968). We will merely suggest a small number of works that will prove helpful to the political scientist who wishes an introduction to the field.

To construct an opinion item is to assume, implicitly or explicitly, a theory of attitude measurement. Three recent collections of readings present selections from the classics of attitude theory and measurement: *Readings in Attitude Theory and Measurement* (Fishbein, 1967); *Attitude Measurement* (Summers, 1970); and *Scaling: A Sourcebook for Behavioral Scientists* (Maranell, 1974). The following works also clarify the links between methods of measurement and scaling and the attitude theories the methods assume: *Techniques of Attitude Scale Construction* (Edwards, 1957); *Theory and Methods of Scaling*

(Torgerson, 1958); *A Theory of Data* (Coombs, 1964); and *Political Research* (Leege and Francis, 1974).

Equally important are several compendia of scales describing measures for almost every concept from anomie to xenophobia. By the use of existing measures, one can take advantage of others' work on the validity and reliability of measures and contribute to the comparability of findings. Bonjean, Hill, and McLemore's *Sociological Measurement* (1967), Shaw and Wright's *Scales for the Measurement of Attitudes* (1967), and Miller's *Handbook of Research Design and Social Measurement* (1970) are useful collections, as are three other works produced under the auspices of the Institute for Social Research of the University of Michigan: *Measures of Political Attitudes* (Robinson, Rusk, and Head, 1968), *Measures of Social Psychological Attitudes* (Robinson and Shaver, 1969), and *Measures of Occupational Attitudes and Occupational Characteristics* (Robinson, Athanasiou, and Head, 1969).

Finally, we should note the work of a Social Science Research Council committee, whose mandate it is to encourage the standardization of background items in surveys. The SSRC Center for Coordination of Research on Social Indicators has convened a working committee for the purpose, chaired by Philip Converse, which has reached agreement on a model set of background items and coding procedures for a number of items including (in its preliminary version) age, sex, marital status, color-race, ethnicity, religion, education, employment status, occupation, income, political party preference, and residential characteristics (SSRC, *Social Indicators Newsletter*, May 1974, p. 2). The widespread use of such standardized items promises great improvements in the precision of estimates of the reliability of background measures.

SURVEY RESEARCH: FOR WHAT PURPOSES?

We have concluded our examination of the types of belief and behavior that surveys typically explore; we turn now to the varied purposes that surveys serve. Without ambiguity we can easily distinguish four very different uses of surveys: basic research, public policy analysis, political electioneering, and commercial profit. In our discussion we will give examples of surveys that were inspired by each goal. Our principal argument, however, is that there has been a fortunate compatibility among the research, policy, and commercial uses of surveys. Perhaps among all the techniques commonly used in social science, none approaches survey research in the degree to which the development of the method is simultaneously tied to each of these different needs.

The Fusion of Research and Policy Interests

The history of survey methodology illustrates the close connection between the research and the policy interests of those who pioneered the development of

surveys. The distinctive policy concerns of the innovators are evidenced in each of four major periods of the method's growth.

Period 1: The social survey movement. In an earlier discussion we observed that the origins of modern survey research are the social surveys of the nineteenth century. It was in these surveys that respondents, the urban poor, were first selected on a sampling basis (Parten, 1950, p. 9). The immediate goal of the surveys was an accurate description of the conditions of the poor in major industrial cities. As we noted, however, this knowledge was to serve different purposes. Some investigators were clearly concerned with stimulating social legislation that would ameliorate conditions for the urban poor. Others seemed primarily concerned with the threat to social order posed by a discontented lower class. Our only point is to restate that these initial social surveys were motivated by the policy concerns of the researchers.

Period 2: The Depression years. Just as the plight of the urban poor led to the nineteenth century social surveys, the depression of the 1930s created an immediate need for accurate data on the social and economic problems caused by massive unemployment. The survey technique was an obvious means of collecting the data, and numerous studies, supported both privately and publicly, were commissioned to study the depression's consequences for American life. As we shall see, the opportunity to investigate these problems attracted the attention of people who were to make, in the course of their data collections, important contributions to survey methodology.

1. United States federal agencies were one such stimulus to substantive and methodological research in the 1930s. Stephan (1948) presents a comprehensive list of the statistical surveys sponsored by federal agencies during the New Deal. Many provided detailed information where none previously existed. Limiting ourselves simply to the *national* surveys, the list includes: The Financial Survey of Urban Housing (1934), The Study of Consumer Purchases (1935-1936), The National Health Survey of 1935-1936, The Continuous Work History Sample, and The 1937 Enumerative Census of Unemployment.

As Stephan observes (1948, p. 27), these surveys clearly reflect their origins in the depression. Beyond the need for data on depression life, the surveys were conducted by personnel who were hired as part of the federal program to create public jobs for unemployed white-collar workers. As a consequence of this last fact, these surveys tended to be crude. The costs of supervising relief personnel discouraged intensive analysis and limited survey designs to those that could be executed by people with relatively little training. Thus, these surveys, as significant as they were, did not usually take advantage of the more efficient sampling procedures that were being developed at that time in several federal agencies.

Even though many of these depression-period surveys did make use of unreliable or inefficient sampling designs, researchers in several federal agencies were at work formulating the complex, multistage sample designs that have since become the recommended procedures for minimizing sampling error and cost. Beginning in 1936 the Research Division of the Bureau of the Census began a study of sampling procedures that led Frederick Stephan, William Deming, and Morris Hansen to develop the five-percent sample in time for use in the 1940 census. In addition to Hansen, Lester Frankel and William Hurwitz were among the important sampling theoreticians actively involved in the Bureau of the Census in the late 1930s and early 1940s.

The Works Progress Administration (WPA) was itself an active research site on sampling problems in the depression years. Lester Frankel was Chief of the Sampling Section of the WPA from 1936 to 1942. The Research Division of the WPA investigated methods of conducting surveys of unemployment, initiating a monthly nationwide survey in 1939. Enlarged and transferred to the Bureau of the Census in 1942, it developed into the Monthly Survey of the Labor Force (Stephan, 1948).

2. Private foundations were a second stimulus to methodological research, working in concert with federal agencies to refine survey techniques and to generate survey data so essential to the policy needs of depression agencies. Indeed, during this period the cooperation of public and private organizations was so extensive that the distinction between public and private sponsorship of research almost loses its meaning. Foremost among the private agencies fostering research on public problems was the Social Science Research Council, incorporated in 1924 and composed of constituent organizations representing American professional associations of social scientists and statisticians.

Throughout the 1930s the SSRC supported committees that focused on depression policy problems. One, The Committee on Government Statistics and Information Services (cosponsored by the American Statistical Association), conducted a thorough analysis of the adequacy of federal statistics and recommended procedures for the coordination of data collection through the auspices of the Central Statistical Board, which had been established by Executive Order on 27 July 1933, "to effect coordination of the statistical services . . . incident to the purposes of the National Industrial Recovery Act" (Social Science Research Council, 1937, p. 8). Stuart A. Rice, perhaps the "father" of quantitative techniques in political science, was acting chairman of the SSRC committee in 1933.

Still another SSRC committee, The Committee on Studies in Social Aspects of the Depression, produced under the direction of Samuel A. Stouffer a remarkable series of 13 monographs, including a study of depression family life by Stouffer and Lazarsfeld (1937) and an analysis of relief policies by White and White (1937). Of these, Stouffer and Lazarsfeld's study of family life is distinctive for its summary of existing social survey data from several countries.

Another example of the symbiosis of the public policy needs and private research efforts is a series of four works on Negro youth in the depression, sponsored by the American Youth Commission of the American Council on Education. St. Clair Drake sets these studies in context:

Hundreds of thousands of American youth were unemployed, and the number of Negroes among them was greatly out of the expected proportion. The New Deal had established a National Youth Administration (NYA) and a Civilian Conservation Corps (CCC) to provide emergency employment, but both educators and government officials were concerned over the long-term welfare of the nation's young people. (About 200,000 Negro youth served in the CCC between 1933 and 1939.) The American Council on Education organized an American Youth Commission to conduct research and to recommend action . . . The key question selected for study was, "What are the effects upon the personality development of Negro youth of their membership in a minority racial group?" *

Four volumes compose this series: Allison Davis and John Dollard's *Children of Bondage: The Personality Development of Negro Youth in the Urban South* (1964; orig. pub., 1940), E. Franklin Frazier's *Negro Youth at the Crossways: Their Personality Development in the Middle States* (1940), Charles S. Johnson's *Growing Up in the Black Belt: Negro Youth in the Rural South* (1967; orig. pub. 1941), and W. Lloyd Warner, Buford H. Junker, and Walter A. Adams's *Color and Human Nature: Negro Personality Development in a Northern City* (1941). In addition to these volumes, all of which are based on original research, Ira DeA. Reid prepared a summary volume on previously available knowledge of Negro youth, *In a Minor Key: Negro Youth in Story and Fact* (1940). Donald Sutherland summarized the major findings of the series and recommended a program of educational and social planning in *Color, Class, and Personality* (1942).

Two aspects of this series attract the interests of the historian of surveys. First, it provides some of the earliest survey evidence on the life and attitudes of American blacks and thus stands as an important baseline for measures of change. Of particular interest is Johnson's study of black youth in the South. As Drake notes in his introduction to *Growing Up in the Black Belt*, Fisk University (where Johnson taught) "had the precise data at hand needed for scientific sampling within the Black Belt, something no other institution except the University of North Carolina had" (Johnson, 1967, p. xiii). Johnson selected eight counties as sampling sites, including two with traditional plantation agriculture, two in declining plantation areas, one near the city of Memphis,

* Reprinted by permission of Schocken Books Inc. from the Introduction by St. Claire Drake to *Growing Up in the Black Belt* by Charles S. Johnson. Copyright © 1967 by Schocken Books Inc.

and another in which Negroes competed with white farmers in a nonplantation setting. In school classrooms, more than 2000 youths completed several highly structured questionnaires measuring personal attitudes, attitudes toward race, and I.Q. [As Johnson states (1967, p. 334), "I.Q. is used in this study as an indicator of the effects of cultural and educational differences between children assumed to be essentially the same (as groups) in inherent capacity."] These tests were followed by intensive interviews with about 20 percent of the sample and with interviews of parents in 916 families. In short, Johnson's survey was an early example of a complex and sophisticated design.

The other books in this series are also important works, but they are not, strictly speaking, surveys. Davis and Dollard's *Children of Bondage* (1964) is based on intensive case interviews of 30 adolescents interviewed several times a week over a period of four to seven months, supplemented with additional interviews with other black adolescents for insights into methods of child-rearing and attitudes toward whites. Similarly, Frazier's study (1940) is, in large part, a case study approach based on neither structured schedules nor rigorous sampling procedures. Nonetheless, though they lacked the trappings of surveys, these latter works provided informative, early insights into the attitudes of Negro youth in the United States. In fact, the potential of these early studies as measures of change did not escape notice. Almost 20 years after Davis and Dollard had intensively interviewed their sample of black adolescents, a research team returned to study these same individuals, who were by then in their middle age. As Rohrer and Edmonson describe the goal of their work, *The Eighth Generation*:

We wished not only to discover what had happened to the "Children of Bondage," but also to find out whether they were perpetuating in their relations with their own children the methods by which they themselves had been trained. (Rohrer and Edmonson, 1960, p. 5)

Second, the books in this series are notable for the opportunities they afforded black scholars. Johnson, Frazier, Adams, and Davis are black scholars who produced numerous books before and after those in this series; their staffs were largely black as well. Johnson built Fisk University into an important center for the study of Negro life and later served as its president.

In sum, the depression years in the United States offer a wealth of examples for our argument that the methodology of survey research developed in a context in which policy needs and basic research were knit together. This is perhaps most clearly illustrated in the important advances that were made in sampling theory by scholars working within New Deal action agencies and the Bureau of the Census. The connection of policy and research goals is evident, as well, in the number of government surveys that provided basic data on the impact of the depression on American life. Finally, we note close cooperation between government and private organizations such as the Social Science Research Council and the American Youth Commission that con-

tributed so much to illuminating aspects of American society about which we had previously been so ignorant. Yet, as distinctive as we might regard the depression years in these respects, the onset of World War II simply presented a new category of policy problems for public officials. As we shall see, survey methodology flourished in the next decade, again largely due to the efforts of social scientists working within federal agencies.

Period 3: World war and cold war. As Sheatsley observes (1963), the official use of United States public opinion data began in 1939 with the establishment of the Division of Program Surveys in the Department of Agriculture. Under the direction of Rensis Likert, the division became a center for the development of survey research techniques. For example, under the tutelage of J. Steven Stock, the division developed the Master Sample of Agriculture in 1943, which divided the rural United States into a grid of small sampling units making possible the efficient, repeated selection of many samples from this single sampling frame (Kish, 1965b, p. 478). Likert himself became a central figure in a controversy over his advocacy of the more frequent use of fixed question-free answer items and of the use of batteries of items to form scales and indexes to avoid reliance on single questions. In an article with the self-explanatory title, "The Controversy over Detailed Interviews—An Offer for Negotiation," Lazarsfeld (1944) made the now conventional argument that both closed and open questions have their appropriate place in interview schedules.

After the bombing of Pearl Harbor, other federal agencies began to call on the expertise of the Division of Program Surveys. Dorwin Cartwright, then a staff member, conducted research for the Treasury Department on people's motives for buying and redeeming war bonds, information that was critical to guiding policy on war and postwar inflation control (Cartwright, 1950, p. 55). On the basis of its survey, Cartwright predicted that the rate of bond redemption in 1946 would fall between \$4.5 and \$6.3 billion and was rewarded with an actual result almost in the center of the interval, \$5.4 billion.

Cartwright's study of bond redemption was merely one of many studies of civilian attitudes and behavior conducted by the Division of Program Surveys. But the Division was not first into the field. D. Caradog Jones (n.d.) reminds us that military purposes prompted the first great English survey—in this case, by William the Conqueror in 1086. The survey is known as Domesday "because it was accepted as a faithful record of facts as they were when it was compiled, and no appeal was allowed against its witness in a court of law" (Jones, p. 15). Jones makes the important point that governments making war must know the resources of their citizens.

But the main purpose of nearly all, if not all, nationwide surveys, undertaken by official direction in different parts of the world in early times, has been to ascertain man-power for waging war and to make an assessment of other resources, in cash or kind, available to fill the coffers of the ruling Sovereign and his Government or those dependent upon them.