

ADVANCING STANDARDS IN AN ADVANCING FIELD  
Paul B. Sheatsley, National Opinion Research Center

As the title of this small paper indicates, I mean to talk briefly about the problem of standards in survey research. I suppose standards can be defined or viewed in at least three ways. First, standards in the sense of standardization or comparability. Certainly it would be desirable if surveys could better be compared with one another, if we all asked certain basic questions such as occupation or income in an agreed upon way, if we all calculated our response rates in exactly the same manner. Second, there are what we might call ethical standards, such as respect for respondent anonymity and willingness to disclose full details of methodology. And third, there are technical standards which might cover such matters as sample size, questionnaire construction, level of interviewer training and supervision, and similar steps in conducting the usual survey.

A very brief history of survey research may be instructive. We must remember that it started in the business world. It was not until after World War II that the campuses took much notice of it. Businessmen and advertising agencies were questioning samples of readers and consumers almost fifty years ago. The Gallup Poll, the first of its kind, was begun by George Gallup, himself a market researcher, and the other preeminent pollsters, such as Elmo Roper, Archibald Crossley and Louis Harris, have also been marketing researchers. That early research 30 or 40 years ago was remarkably simple, even simplistic, by modern standards. They selected their samples by setting various kinds of quotas, they asked one or two questions where today we would spend five or ten minutes on that one topic, and the low budgets and competitive pressures would have left little resources for interviewer training and supervision. Secrecy of exact methods used was commonplace, and it may be presumed that many first who hastened to set up their own survey shops in those days often engaged in fairly sharp practices.

With World War II survey research came of age. Men like Paul Lazarsfeld, Samuel Stouffer, Hadley Cantril and Rensis Likert lent their services to the government in the organization and direction of survey research activities. Such advances as probability sampling and Guttman scaling derived from this government work. More important, a whole second generation of research men who had worked under the "giants" I have named went back after the war to campuses or to commercial activities, and continued to teach and practice the art of survey research. Within two years after the way there had come into existence an American Association for Public Opinion Research, one of whose stated purposes was to improve standards. Leaders in the field had by that time concluded that the old ways of secrecy and every man for himself which prevailed in an earlier day were inappropriate for what had become a mature field of practice considerably affecting the

public interest. AAPOR, as it is commonly called, has always been open to any individual with an interest in public opinion research, and while the majority of members are from the world of business or market research, more than one-third are from universities, government agencies, foundations and other not-for-profit institutions.

AAPOR in its early days almost foundered over the issue of standards. Some academic and government researchers considered some market researchers as rather shady types who ought to be made to shape up. Some market researchers considered the academics as ivory-tower types who were about to impose upon them a set of idealistic standards which could not possibly be fulfilled in the market place. Fortunately, cooler heads prevailed and AAPOR has grown increasingly strong and influential. But it took the Association several years to get around to discussing standards again and when it did it emerged, after another period of years, with a Code of Professional Ethics and Practices, with most of the emphasis on ethics. The American Marketing Association went through a similar experience, and it too now has a Code of Ethics.

These codes define the researcher's obligations to his respondents, to his client or sponsor, and to the public. To these latter two groups, the essential obligation can be summarized as full disclosure. No relevant information, either of substance or of method, may be withheld. In other words, the codes does not dictate the number of cases the researcher should collect on a given study, but it says he must disclose in his report the size of his sample; he is not told what is a satisfactory completion rate in a mail survey, but he is obligated to reveal the completion rate on which his data are based. This of course places the onus for getting good research, and perhaps properly, on the user of the survey, who is presumed to be sufficiently sophisticated, or to have access to sufficiently sophisticated advisers, to recognize loaded questions, faulty sample designs, and so on.

Caveat emptor.

While I can only speak for AAPOR, and not AMA, I can testify that the code has had some measure of success. Every AAPOR member attests that he will abide by the code when he joins the Association, and there is reason to believe that violators are rare. Non-members of AAPOR, of course, are not so bound, but even here there have been successes. AAPOR has an active Standards Committee and any member with a complaint can present the facts to that committee. The committee has no real enforcement power, but it has not yet been pushed to its ultimate sanction: publication of names of offenders in AAPOR's journal, the Public Opinion Quarterly. In one important case, a non-member was making large numbers of telephone calls in cities all

over the country, ostensibly conducting a survey to find out the extent of the market for new automobiles. Actually, this was a cheap and easy way of compiling a list of prospects, and respondents who had indicated any intent to purchase, found themselves confronted a few days later by an aggressive salesman. This was an obvious violation to respondent anonymity in a purported survey and, if continued, could seriously affect public trust in all surveys. When a few members of AAPOR's Standards Committee arranged a visit with the offending researcher and discussed the matter with him, the practice was stopped. There have been many similar instances.

But the caveat emptor philosophy has certain limitations when there really is no emptor, or buyer, but when the survey is done for selfish or partisan purposes and the results fed directly to the public. The man in the street is hardly capable of recognizing poor sample design, bad question wording, or out-of-context interpretation. We have probably all been exposed to examples of private polls which purport to show that a particular political candidate is way ahead in a particular state, or mail surveys which indicate that large majorities wish to take some particular kind of political action. The professional political pollsters such as Gallup, Harris, Crossley, Mervin Field in California, and others, have been seriously concerned by the appearance of numerous "private surveys" in any major election campaign, and they have given much thought to how to impose "standards" in this area. Their response was much the same as that of AAPOR and of the American Marketing Association.

They set up a Council on Published Polls, which includes all of the major polltakers who regularly cover election contests, and they drafted some standards of disclosure which they have circulated to Congressmen, state and local officials, newspaper editors and other media representatives. The standards suggest questions which readers of the purported survey data should ask, and all members of the Council have pledged that they themselves will frankly answer all such questions: The questions themselves cover the usual areas: size and design of the sample, completion rate, question wording, means of data collection, dates of interviewing, and so on. Certainly such simple educational measures cannot stop the appearance of fraudulent or misleading polls, but their existence cannot help but have a long-run effect as news editors become better informed about polls. Meanwhile, the very existence of the Council and its standards makes it much easier to discredit an inadequate poll. The guilty polltaker, instead of seeing his data widely published and accepted, is forced on the defensive when he cannot or will not answer legitimate questions about his method, and in many cases his releases may be thrown away without ever seeing publication at all.

In these various ways, then, professional associations have tried to grapple with the

problem of standards. But what is the present situation? Surveys and polls are more popular than ever. Federal and state governments want "evaluations" of particular health, educational and welfare programs which can only be conducted by means of survey research. Business firms more and more rely on survey research data for marketing decisions. Often the sponsors of the surveys have little or no understanding of methodological problems and they are always highly resistant to large expenditures of cost and time. The pressures for cutting corner here and there inevitably lead to a lowering of standards.

The popularity and apparent simplicity of survey research even attracts what we might call the "innocent ignorant". A friend of mine, an eminent psychologist who has been working with survey data most of his professional career, was recently appalled to learn that a chemistry professor on his campus had sent his students into various local supermarkets to question housewives about their attitudes and knowledge concerning the presence of phosphates in detergents. The good professor had no knowledge of how to obtain a truly representative sample of housewives, or even whether housewives were really his intended universe; he had no experience with questionnaire construction; his students had not a glimmer of interviewing problems and techniques. His survey, besides giving the students a faulty understanding of survey research, probably antagonized some of his respondents who had already been approached by other amateur surveyors and who may have been put off by the inept questioning and approach of the student interviewers, and in the end this survey produced data which may have been quite seriously misleading. Similar surveys are doubtless conducted every day by high schoolers, local do-good societies, small business, and of course newspapers.

It is clear that the problem of advancing standards in the field of survey research is not a simple one. We can probably not insist that that chemistry professor be required to take courses in survey research and pass an examination, any more than we can demand the same for the high school students, the members of the local Women's Club, or the small businessman. What AAPOR has tried to do, in at least a preliminary attempt to solve the problem, is to make expert advice available to such groups. The AAPOR membership, while concentrated heavily in New York, Washington and other major research centers, is scattered widely over many cities and college campuses. Members are urged to be alert for opportunities to be of local professional service; inquiries direct to AAPOR are referred to a local member, or the names of several members are provided to the inquirer; mailings have gone out to college campuses advising them of AAPOR's desire to be of help in consultation on survey research matters. Individual members sometimes establish formal consultant relationships with local groups, or more often simply donate their advisory services as a professional obligation.

You will have noted that all of the efforts I have referred to, to establish or maintain standards, have dealt with ethics rather than techniques, and have been of an educational rather than punitive variety. Such efforts are necessarily slow and imperfect. One may ask, why cannot categorical rules be set down for survey research standards and those standards rigidly enforced? It should not be too difficult to set reasonable standards for sampling, completion rates, question wording, interviewer training, and so on. We can all recognize abuses when we see them. But, in fact, a moment's thought, or even less, reveals very clearly that operational standards for survey research are impossible to set.

We sometimes call our work a science, and indeed a good piece of survey research will meet the scientific requirement of being capable of replication by an independent researcher following the same methods. But it is obvious that certain areas of survey research are still very much of an art, and particularly so when we are dealing with measures of attitude, future intentions, or beliefs, rather than factual or behavioral information. There is no one correct way to write a survey question, nor even any way to know exactly which question or questions should be asked. In spite of a great deal of empirical and experimental research, there are no hard and fast rules which govern the selection of interviewers, nor is the same method of training and supervision appropriate for all of those hired. Even assuming such standards were possible and maintained, there is no way a researcher can prevent his data from being quoted out of context or manipulated in improper ways, so that all his methodological precautions come to naught.

Aside from the fact that we simply do not know enough to dictate standards in many areas of survey research, there is the fact that standards simply have to be relative to the time and cost resources available. The cost of a survey - a survey - can range from a few hundred dollars to several hundred thousand dollars. Obviously, if one has unlimited time and money, one can design and perfect a survey with extremely high standards. While our resources are hardly unlimited, we at NORC have had recent experience with an ambitious evaluation study of federal manpower training programs, sponsored by OEO and the Department of Labor. Enrollees in five programs in ten cities have been interviewed four times over a two-year period; this sample has been matched with a group of controls, selected from house-to-house screening, who have been similarly followed up. The demand for a final completion rate of 80% in this mobile and hard-to-find sample has required the setting up of separate NORC offices in each of the ten cities; rigorous training and supervision of an appropriate interviewing staff; and the expenditure of an enormous amount of time, effort and money to locate and interview reluctant or elusive individuals.

But the point is that such expenditures are hardly required on most surveys. The marketing man who wants to know which of two advertisements is more attractive to the public, or the community agency who wants to know how exercised the public is about environmental pollution, can make do with much less. Furthermore, the researcher rarely if ever has unlimited time and cost resources. He has a deadline and he must stay within a particular budget, whether it be generous or miniscule. In effect, then, he has to design his survey within those constraints. He can't afford a full probability sample, so he uses quotas or some means of weighting for persons not at home. He can afford only a half-hour interview, so he throws away all the batteries of questions he would like to ask about personality characteristics. He can't afford a full probability sample, so he uses quotas or some means of weighting for persons not at home. He can afford only a half-hour interview, so he throws away all the batteries of questions he would like to ask about personality characteristics. He can't afford to supervise the interviewers as well as he would like, so he pretty much accepts what they give him. His report is due in three weeks, so he can't pursue all the lines of analysis he had intended. All this is not necessarily bad. A survey can be over-designed and too perfect for the job intended; some reasonably accurate information is usually better than none at all. But it sure makes the job of setting standards difficult.

Finally, beyond the fact that we do not know enough about some areas and that the same standards are not appropriate to all surveys, any set of standards must always be based on past experience and present knowledge. The title of this paper refers to an "advancing field". Survey research has advanced tremendously in my own time, and it continues to advance. When Gallup first sent a national staff of interviewers out to select a representative sample by the use of sex-age-economic level quotas, this represented a tremendous advance over the Literary Digest's use of mail ballots. The growing demand for probability sampling represented a similar advance over quota sampling. The contributions of numerous individuals and agencies have improved our knowledge of all phases of survey research, and the introduction of the computer has of course revolutionized our means of processing survey data. We are perhaps on the verge of even more exciting discoveries.

The techniques of survey research are being employed in ever more sophisticated ways. Merely within the past six months, we at NORC have been faced with the following problems of research design and execution.

Personal interviews with all practicing internists in a defined suburban area. Since many of the physicians knew one another, the problem was to avoid contamination effects.

Identification of the population of working journalists in the United States and personal interviews with a representative sample of these.

Establishing a panel of patients suffering from Parkinson's disease and following these up over a period. Half of the group are receiving a new method of treatment, half are controls.

Sampling patients served by the emergency room of a large metropolitan hospital and following up several weeks later with interviews in their homes.

Validation of elementary school data submitted to the Office of Education, through interviewer inspection of records at a sample of schools and brief interviews with school officials and parents.

The list could go on and on. The point is that each of these assignments involved a

great deal of fresh thought on such problems as defining the universe, inventing an appropriate method of sampling, gaining access to the intended respondents, and obtaining valid data from them. And naturally there were the usual constraints of time and cost in each case. Under these circumstances, no conceivable set of general operational standards for survey research could offer much guidance.

In sum, then, I hold that there has been and will continue to be advancing standards for survey research. In almost every aspect of research, our workmanship today is far superior to what it was twenty years ago, or even ten years ago. But these advancing standards have come about not through any codification of standards approved by some official group, but through the inventiveness and increasing sophistication of the sponsors, practitioners and users of survey research. There will always be abuses of the survey method, as of many other things in this imperfect world, but it is becoming more and more difficult for such surveys to gain any measure of acceptance. More than this, we can probably not reasonably expect.