

Close Enough for All Practical Purposes

REUBEN COHEN

JUST two days before our meetings began at Buck Hill Falls an article in the *New York Times* attracted my attention—"Questionnaire sparks tumult in academe."¹ Well, I thought, this sounds like a good warmup for our AAPOR Conference. It turned out that this particular discussion has been going on for a couple of years—long enough, according to the article, for a file of some 515 pages of articles, letters, and other documents to be accumulated. I have not seen "the file," and I do not propose to add to the tumult. But the fundamental issue posed by the dispute is one of quality.

Here are some sample comments from the article:

Getting things right in social science research is not easy. [I believe that one.]

The sample of potential respondents was a "hotch-potch of various procedures."

I wouldn't trust any survey with a response rate like that.

[The authors] . . . were not adequately careful according to the standards of their own discipline.

The last remark is one that should concern all of us, and it has been one of the central themes of this conference. Who sets the standards? How should our work be evaluated? Who can tell the good from the bad?

Some 30 years ago—plus or minus 5 percent—Jack Elinson handed me a reprint of W. Edwards Deming's list of errors in surveys.²

¹ Edward B. Fiske, *New York Times*, May 29, 1979.

² Deming, *Some Theory of Sampling* (New York: Wiley, 1950).

Reuben Cohen, Response Analysis Corporation, was President of AAPOR in 1978-79. This is an adapted version of his presidential remarks at AAPOR's 34th Annual Conference in Buck Hill Falls, Pennsylvania.

Deming's list of possible errors is a classic. I still take it out and look at it every once in a while. It lists 19 categories of errors in surveys. If you look closely, you will see that most categories are further divided into specific types. When I counted them recently, I found that the list contained more than 60 different types of errors.

Of course, the message was pretty obvious: Now that you know about them, don't make them. With my relative inexperience, and my eternal optimism, I accepted the challenge. My first approach was to try to do the perfect survey. I am still trying, but I should know better. I quickly discovered Murphy's law—if anything can go wrong, it probably will. But I also discovered something else. Even without the time and budget constraints that most of us complain about, there are no perfect surveys. Every survey has its imperfections. The world is not ideally suited to our work. The best we can do is think through the ideal approach to a survey design, or implementation, or analysis problem—what would we do if we had our druthers?—then get as close to the ideal as we can within the constraints of time and budget which govern much of our work.

My purpose in these remarks is not to suggest to you that errors exist in surveys. Everyone here over the age of eighteen has probably seen them all. A number of them have been discussed in sessions of this conference. The problem has been, is now, and will continue to be—how do we detect errors, control them, evaluate them, compensate for them when possible, and in general determine the effects on survey findings. How do we really tell the good from the bad? Equally important for those of us who do surveys, or who interpret or report their findings: How do we fulfill our responsibilities to our sponsors and other publics to communicate the limitations of our findings, and to display or at least suggest the possible imperfections in survey procedures and results?

Where do we turn for guidance in our day-to-day work? I think there are some helpful signs. Just over the past few years the pace has accelerated. There have been new insights into old methodological problems, new developments in technical standards, and additional pressures for disclosure of at least a minimum amount of information about how the work was done for published survey results. The increased activity is coming from a variety of directions—from professional associations, from government agencies for government-sponsored programs and surveys, and from business organizations.

AAPOR, of course, has had a code of professional ethics and reporting standards for more than 20 years, but it was only four years ago that procedures for dealing with alleged violations of the code were adopted and first applied. The procedures are controversial to

be sure, they require much time and effort to carry out, they can probably be improved, but they are a start in the right direction.

A number of other organizations have joined with AAPOR in developing and publicizing a variety of standards. For example, the recently created National Council on Public Polls has established standards for information to be released about published or broadcast reports of polls.

The Council of American Survey Research Organizations (CASRO) was formed a few years ago, and adopted a code of standards just last year. The Market Research Association and American Marketing Association have also adopted codes, and a set of standards is currently being considered by the Survey Research Section of the American Statistical Association.

Four years ago the Bureau of the Census published a set of standards for discussion and presentation of errors in survey data.³ These standards were designed to better inform data users of the important limitations of the Bureau's estimates, limitations due both to sampling and to response and other nonsampling errors.

Other major undertakings have been underway by the Committee on National Statistics of the National Research Council; some of its work has been published, and we can look forward soon to seeing the results of its study of practices in the handling of incomplete data in surveys. Standards for surveys published by such companies as General Foods and Sears—and I'm sure there are others—are another source of guidance for us.

Given the proliferation of survey codes and standards, and the major methodological work now underway, just how good are survey practices? A year ago, a report of a pilot survey of surveys was published by the American Statistical Association. The pilot survey was supported by the National Science Foundation and directed by Barbara Bailar and C. Michael Lanphier.⁴ It included an evaluation of 36 surveys, most but not all sponsored by the federal government. It was not intended to be definitive, and the authors caution against generalization beyond the 36 surveys. But for the surveys they studied the authors concluded that:

—Twenty-two of the 36 surveys did not meet their objectives, either because of poor design or serious technical flaws.

³ Marie E. Gonzalez, Jack L. Ogus, Gary Shapiro, and Benjamin J. Tepping, "Standards for discussion and presentation of errors in survey and census data," *Journal of the American Statistical Association* 70:351 (1975), Part II.

⁴ Bailar and Lanphier, *Development of Survey Methods to Assess Survey Practices*, (Washington, D.C.: American Statistical Association, 1978).

- Samples were, for the most part, poorly designed; very little documentation existed.
- Survey response rates were difficult to collect and compare.
- Quality control over data preparation and processing activities varied considerably.

And so on.

At this point it is not certain that the survey of surveys will be extended beyond the pilot stage. I hope it is. But the results already reported by Bailar and Lanphier are instructive and important. We need to know more about actual practices in survey research. We need more, not fewer, technical standards. As much as we may complain about outside review agencies looking over our shoulders, or delaying our work, they perform an important function. I am thinking of such technical reviews as those performed by the Advertising Research Foundation for media evaluation studies, and the Office of Management and Budget for federal government-supported surveys.

As a group, the people gathered here—and our associates minding the stores back home—do good work. Let me share with you a personal prescription for making it even better. The next time you decide on a particular approach to a survey problem, whether in the design or analysis, take time to work out at least one alternative. Chances are the alternative will appear to be more difficult, to take more time, to be more costly. Examine carefully those assumptions.

If it appears to be more difficult, who knows, you may be able to get other people to do the work for you. If it appears that it will take more time, ask whether the additional time will really make very much difference. If it appears that it will be more costly, try to figure out what the difference would really be. After all, there is just one true measure of survey quality. It is the amount of useful information delivered per dollar of survey costs.

But the other possibility is that you will be surprised, as I have been on a number of occasions, to find that on careful appraisal of alternative methods, the preferred procedure is not more difficult, or more time-consuming, or more costly. It's just different from what we have done in the past.

Les Kish has suggested in regard to sampling that "practical work consists in good part of guessing what irregularities, where, and how much, one can afford to tolerate."⁵ The same is true of other aspects of our work. It should be done well. It can and should conform well, even if not perfectly, to an ideal approach. To illustrate the principle

⁵ Kish, *Survey Sampling* (New York: Wiley, 1965).

involved, let me conclude with a story that I have also seen attributed to Les Kish. It seems that a beautiful woman had two suitors, and could not decide between them. One was a mathematician, the other a statistician. Because she was wise, as well as beautiful, our heroine devised a contest that she thought would be fair to both. She placed herself in the center of a square room, directed her two suitors to opposite corners of the room, and proceeded to explain the rules of the contest. "I will flip a coin to decide which one of you will go first. The rule is that you may advance half the distance toward me. You should then take turns, using the same rule each time. Each time you take your turn you may advance half of the remaining distance toward me." At that point, the mathematician abruptly turned and left the room, muttering that the contest could have no winner since neither he nor his rival could ever reach the woman each of them loved. The statistician smiled and stayed on, saying, "I can get close enough for all practical purposes."