PRESIDENTIAL ADDRESS
A RESPONSE TO THE NONRESPONSE PROBLEM

NORMAN M. BRADBURN

One day during the French Revolution, three prisoners were led to the guillotine—a politician, a journalist, and a scientist. As the politician was put into the guillotine, he was asked if he wanted to go face down or face up. Since politicians never like to face a problem until they have to, he chose to go face down. When the rope was pulled to let the blade fall, the mechanism jammed and the blade stuck. After some messing around, the executioners decided that the sentence had in principle been carried out, and the politician was free to go with his head in place. Next the journalist was led up. He was asked whether he wanted to go face down or face up. Like all good journalists who can't stand not to know what is coming, he chose to go face up. Again the rope was pulled, the mechanism jammed and the blade stuck. After some more messing around, the executioners again decided that the sentence had in principle been carried out and that the journalist could go free.

Finally the scientist was led up to the guillotine and asked whether he wanted to go face down or face up. He thought about it a minute and decided it was better to face the problem squarely, so he also chose to go face up. The rope was again pulled and again the mechanism jammed and the blade did not come crashing down. As the executioners were about to give up, the scientist said, "Ah, I see the problem; you just need to adjust that weight on the left there and it will slide down fine." I will leave it to you to decide what came next.

I want to discuss a topic today that many of us will feel is putting us uncomfortably close to the scientist in this story—that is, the question of response rates in surveys. My thesis is that we must not take the path of the politician and turn our face away from the problem, nor can we afford to take the path of the journalist, who may or may not have seen the problem but did not say anything about it when his own head was at stake. Instead, we must face the problem straight on, even if we risk endangering our livelihoods.

Public Opinion Quarterly Volume 56:391–397 © 1992 by the American Association for Public Opinion Research
 0033-362X/92/5603-0026$02.50

The Nonresponse Problem

We all believe strongly that response rates are declining and have been declining for some time. Part of the problem is locating respondents, and part of the problem is getting cooperation. Locating respondents is more difficult than it used to be because of life-style changes, especially among women. People just are not at home as much as they used to be. They also appear to be harder to convince that they should participate in a survey.

Hard data are very difficult to come by. The annual Walker Industry Image Survey reported that 36 percent of their respondents (who were themselves only 46 percent of their original sample) said that they had refused to participate in a survey in the past year, up from a low of 15 percent in 1982, but not significantly different from the 34 percent in 1988. A *Washington Post* article a few weeks ago quoted Larry Hugick of Gallup as saying that response rates had dropped from over 80 percent in the mid-1970s to around 60 percent today. It is not clear how much of this decline is due to refusals and how much is due to greater difficulty in finding someone at home or being unable to make contact through an answering machine. On the other hand, the General Social Survey has shown no consistent pattern in response rates since its beginnings in 1972. A few years ago the AAPOR Council appointed a small task force to look into the response rate problem, but they abandoned their efforts because they could not get enough comparable data.

The reasons for the lack of serious concern about response rates, I believe, are complex. I would group them into three main reasons: (1) the competitive market, (2) a belief in the robustness of the survey method, which is a way of saying that it isn't really a serious problem, and (3) a belief that statistical adjustments can take care of the problem. Let me discuss each of the reasons briefly.

The survey research world is a very competitive one, a fact that I do not need to explain to this audience. One of the desirable effects of competition for buyers is that it drives down prices to the lowest level compatible with getting sufficient products produced at the quality level demanded by the market. Response rates are one, but only one, element of quality in surveys. Timeliness is another element that is also often very important. I would interpret the lack of interest in response rates on the part of most of the buyers of commercial survey research as an indicator of the relative lack of interest in this aspect of quality. Consequently, there is a lack of interest in paying the higher costs or waiting the extra time that go with obtaining higher response rates. Where there is a concern for response rates as a quality measure,

as, for example, in the Census Bureau and in much federally sponsored survey research, very high response rates are demanded, and the clients are willing to pay the necessary costs and take the time to obtain them.

Second, this lack of interest in response rates on the part of most of the market is due, I believe, to the belief on the part of both clients and of many survey researchers that the survey method is a very robust one and that the data are not seriously affected by even fairly low response rates. For example, *Science* magazine, the official publication of the American Association for the Advancement of Science, saw fit recently to report the results of a survey they had done on a very politically sensitive topic, academic misconduct, that had a 31 percent response rate. While they noted that the response rate was low and that there might be biases, the editors believed that the survey was still robust enough to be worth reporting, even though I am sure that they would not have accepted an article based on that survey if it had to pass peer review.

Why should we believe the method is so robust? The reason, I believe, is that much of our experience suggests that it is. Let me give you two examples of the type of experiences that would lead researchers to believe that low response rates may not be as great a problem as some would suggest. The first comes from quota sampling. When the predominate mode of data collection was face-to-face personal interviewing, one of the most vigorous debates in the survey world was the degree to which one could relax the requirements of probability sampling and substitute quotas for certain demographic characteristics without undermining the validity of the inferences that could be made from the data. As my colleague Martin Frankel is fond of pointing out, a well-designed quota sample which relaxes probability sampling at only the last stage of sampling is like a full probability sample with a 33 percent response rate, because that is about the percentage of households that are interviewed on the first call. Quota sampling works well for many purposes, and controlled experiments between sampling methods often produce little or no differences in data. The difference, of course, is that probability sampling rests on the mathematical theory of sampling, while quota sampling does so only approximately. Thus with quota sampling we lose the theoretical basis for drawing inferences from the data, and we never know when (or how often) our quotas will lead us astray.

A second example comes from preelection polls conducted by telephone. Preelection polls are valuable to the survey research profession because they are one of the few types of surveys that are public and are fairly frequently done and whose results can be readily compared

with an external validity criterion—the election results. Preelection polls are done with varying amounts of constraints such as time and money and have varying response rates. Yet the final preelection polls are rarely very far off in their predictions, at least by the standards that the clients, including the public at large, hold them to. One could argue that the standard isn't very high since the standard is being right about the winner, not about the accuracy of the actual vote percentage. It is also likely that somewhat more attention is given to response rates for final preelection polls, since those are the ones by which the pollsters will be judged. Although I do not have any data, I would expect that, given the costs of being wrong in that market, serious attention would be given to response rates if it were the case that response rates had a very big effect on the accuracy of the preelection polls.

While I think that such practical experiences are the source of much of our relative lack of concern over response rates, there is another factor that I think contributes to our confidence in the robustness of the method. We all learned in our statistics courses that sampling error is a function of the variability of the phenomena and the size of the sample. The fact that sampling error is a function of sample size and not the size of the population is drummed into every student who takes even the most elementary sampling course.

But sampling theory was developed before population-based sample surveys were very common, so the question of response rates was not of much concern to sampling theorists. Typically, sampling texts do not tell us what happens when the sample is not fully executed. How does the standard error of a sample of 100 with a 100 percent response rate compare with the standard error of a sample of 100 with a 50 percent response rate? There is no theoretically derived relationship between sample size and sampling error which involves a term for response rates. Thus, on the surface, a sample of 100 could be used to calculate the sampling error in the same way regardless of the response rate obtained in getting that sample.

Intuitively, it does not seem a reasonable proposition that response rates have no effect on sampling error. Sampling error, however, is a function of two components, sample size and variability of the phenomena being studied. The measure of variability comes from the sample data themselves since we rarely know the variability of a population independent of the sample data. Low completion rates endanger the accuracy of the estimates of the variability and thus potentially affect the accuracy of the measure of sampling error. Unfortunately, there is no mathematical function that can tell us in advance what the effect of the different response rates will be, and there is no empirical

method for computing the effect of the nonresponse rate on the standard error as there are, for example, methods for computing design effects.

Perhaps the more important effect of response rates, however, is on the potential bias in the sample. While sampling error calculations, if correctly done, tell us something about the "margin of error" in our data, they do not tell us about bias. If samples are correctly drawn, the data should be unbiased, at least from a sampling point of view—they might, of course, be biased from other sources such as the way the questions were worded, the context in which they were embedded, or the way the interviewer asked the questions. Low response rates can produce biased samples, particularly when probability of response is correlated with some important characteristics, as, for example, for preelection polls, probability of voting, party identification, or important demographic characteristics like education, income, race, or gender.

It is this potential bias problem that leads to my third topic—weighting. Studies of nonresponse indicate that nonresponse is not randomly distributed across the population, but tends to be higher among those at both ends of the income distribution, among the elderly, for men, and for those with limited English proficiency, to mention a few of the most important characteristics. Greater callback efforts can raise the response rate by decreasing those who are not reached because they are infrequently at home, and reassignment of cases to highly skilled interviewers can reduce the number of refusals by "converting" some respondents, but demographic biases in the sample are almost certain to continue to exist. Weighting the sample up to known population parameters is believed to reduce the bias in the data because it is assumed that the missing respondents would have given the same type of answers as those of similar demographic characteristics who were contacted and did answer the survey questions.

The potential flaw in this argument is that it assumes that there is no correlation between willingness to respond and the substantive answers to the survey questions. In microcosm, this problem is the same as that which we rage against in what I have called "SLOPs," for "self-selected listener opinion polls." The responders in scientific surveys have a certain element of self-selection, albeit under the very restricted limits of the sample selection. But, nonetheless, they are more cooperative, and the possibility that there is some correlation between their opinions and their cooperativeness is not to be dismissed lightly. Thus, with declining response rates, what is mostly now treated as an interesting possibility will have to be taken more seriously.

What Can Be Done?

I turn now to potential solutions. First I would like to recast the problem. We often talk about response rates as if they are a fixed characteristic of surveys whose movement we can observe. I would frame the issue differently. Within fairly broad limits, response rates are a function of the resources, both time and money, devoted to them. It is possible to get high completion rates if we devote the necessary resources to the task. The problem of declining response rates is better described as a problem of increasing costs to get the same response rate. For most of the survey market, there is an unwillingness to pay the increased costs, either in time or money; thus we observe the declining rates. But we should not view them as intrinsically declining.

If response rates were to become a more salient aspect of survey quality and there were a greater realization that higher response rates are possible at a price, then we might see a rising demand for surveys with higher completion rates. The net result might be fewer surveys, each of which costs somewhat more and has a higher response rate. Such a change might not be an entirely undesirable outcome, but it is not a probable one.

Assuming, however, that clients are not willing to pay more for individual surveys, even if their total survey budgets were unchanged, or wait longer for the results, are there other things that can be done? The answer is "Yes." Resources can be more efficiently used in achieving a higher effective response rate. There are methods which have been advocated for at least 50 years that allocate resources differently according to the difficulty of getting responses.

These approaches require us to alter the strategy we employ most of the time, which is to make the maximum effort we can afford to complete as many of the sample cases as possible, working all cases until the data collection budget is used up. An alternative method is to start data collection with a low-cost method applied to all cases, such as trying all numbers two times. The results from this method, of course, will be relatively low. But instead of proceeding to use a rule such as "Try all remaining cases two (or four or eight) more times," select a sample of the remaining cases and concentrate resources on achieving a very high response rate for those cases, such as making 18 to 20 calls, using specialized "converters," paying significant respondent fees, or perhaps even employing other modes of data collection. When these subsampled cases are then combined with the original sample cases, they will have to be weighted because their probability of selection is different from the cases that were given the inexpensive treatment. These methods yield higher, but weighted,

response rates and reduced bias, although we lose something in terms of effective sample size. A number of these methods are described in the excellent book *Survey Costs and Survey Errors* by Bob Groves.

These methods share one element in common, that is, learning more about the characteristics of the hard-to-get respondents and, with some additional effort, even the nonrespondents. The data on these difficult respondents can then be used to do better weighting that can incorporate more information into the weighting procedure. While we may not have the resources to get high response rates across the board, we can allocate the data collection resources in a more targeted manner to learn more about the possible bias arising from low response rates.

If practical considerations of time and cost make it unlikely that clients are going to alter their requirements to enable us to reduce nonresponse rates, we must look to methods that enable us to get more information about those we miss and use that information in our analyses. The alternative methods just described focus more on understanding the nonresponders and using this information to adjust the data more intelligently. There is considerable room in our practice for increasing our understanding on nonresponse without great increases in cost.

We also need help from statisticians to develop methods that enable us to model empirically the effect of nonresponse on the accuracy of our estimates. With the great increase in computing power that has been recently achieved, it is possible to use calculation-intensive methods for adjusting data that can take into account many variables. There have been advances in the statistical basis for treating such problems as missing data within individual cases. Similar attention needs to be given to treating missing sample cases.

Until we feel comfortable about our understanding of the characteristics of the nonresponders and incorporate this information into our analyses of the data, we cannot feel comfortable about delivering reports to clients based on surveys with low completion rates. Perhaps we cannot do much to increase the response rates, but we can bring more attention and sophistication to the treatment of data with low completion rates. We need to face the problem squarely and solve it, even if there is a risk that the guillotine may come down before we can get our heads out of the way.

Reference

Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: Wiley Interscience.