TOWARD A THEORY OF RESPONSE EFFECTS IN SURVEY RESEARCH

Seymour Sudman, Survey Research Laboratory, University of Illinois

It is a wise experimenter who knows his artifact from his main effect; and wiser still is the researcher who realizes that today's artifact may be tomorrow's independent variable.

William J. McGuire

The importance of response factors on survey results is now widely recognized by survey users and practitioners. There are literally hundreds of studies that demonstrate some kinds of response effects or non-effects. Interest in the problem has been heightened by the work of Rosenthal and others in experimental psychology on the presence of experimenter effects and other artifacts in behavioral research.

As McGuire wittily points out, there are three stages in the life of an artifact.

A. The Ignorance Stage

B. The Stage of Coping

C. The Exploitation State where:

> It is rather heartwarming to observe that in the final stage in the career of an artifact, the variable comes into its own. The Ugly Duckling becoming the Prince Charming which gives rise to a new line of research.

We are now well past the ignorance stage, and are into the stage of coping. Although there have been some examples of exploitation such as the notion of acquiescence or yeasaying, in general, survey research is still quite far from exploiting response factors. One limit to coping and exploitation has been the lack of general theoretical structure of response factors. Most studies have demonstrated response factors in highly specific situations that are difficult to generalize.

In this monograph, we attempt to present a more general framework than has been available up to now. It would be presumptuous to claim that we have developed a general theory of response, but this is the direction of this research, and hopefully the results presented here are the first stage in the development of such a theory.

Methodology

The first stage of the project consisted in preparation of a bibliography of books and articles related to the topic of response effect. This bibliographic search was greatly aided by a more general bibliographic file on survey methods that had been compiled under Charles Cannell's direction at the Survey Research Center, University of Michigan. As of now, the bibliographic file for this project consists of about 1,000 items and new sources are still being added. This bibliography is included in this report in its entirety, although about half the items could not be coded using our coding system.

Throughout this study, the dependent variable remains the relative response effect. Many studies gave information that demonstrated response effects, but where the magnitude and direction of these effects could not be determined. Since we were attempting to measure the effects of a large number of independent variables on response, we needed sufficient data on both the actual figures reported and some validation measure to establish a reasonably accurate estimate of the magnitude of the relative response effect.

The "field work" for this project consisted of obtaining copies of the studies and coding them in a uniform format to be described in detail below. The sample then consisted of the information coded from these studies. The number of items of information from a single survey varied from one to several hundred. Thus, when one uses unsummarized results, these are weighted towards the studies that provided more items of information. In some of the analyses, all results from a given study are combined so that the studies are given equal weight.

Where the study dealt with actual behavior such as purchasing, voting, owning a driver's license or a library card and outside validation sources were available it seems appropriate to call the differences response error or bias. Where the study related to attitudes and there were no outside validation sources, it seems better to use the more neutral phrase "response effect." Since many different types of studies were included, absolute errors become unworkable since there is no way of combining studies. A measure of relative effect was adopted and defined to be:

$$RE = \frac{(\text{Actual-Validating})}{s}$$

where s is the standard deviation of the population, obtained from the validation information, if possible. Where no data were given on the size of s, an estimate of RE was made using $\frac{(\text{Actual-Validating})}{\text{Validating}}$. This estimate is satisfactory for populations where the coefficient of variation is near to one, but becomes quite poor as the validating mean approaches zero. In those cases, the results were omitted.

Where the studies reported about additudinal information and no outside validation was possible, the weighted mean of all observations was used for validation. Thus, for example, in a study which contrasted responses of blacks to black and

white interviewers, the grand mean was found by combining the responses to black interviewers with those to white interviewers weighting by the sample sizes. The relative effect was found by computing the difference between the response given to black (or white) interviewers and the grand mean, with s being computed from the grand mean.

Analysis

The basic (perhaps simple-minded) approach in analyzing the results is to treat the several thousand items of information as one would treat responses to a typical survey. One observes differences in the dependent variable (here response effects) and one proceeds to search for the combination of independent variables that best explain the results. One never finds complete consistency in the real world. Thus, in studies of prejudice, it is generally the case that prejudice by whites against blacks declines with increasing education, but regional factors are also critical. College graduates in the South have higher prejudice scores than do Northern respondents with only an elementary school education. Nevertheless, there are some Southerners with low education who show no prejudice.

Similarly in this analysis of response error, we are unable to find absolute truths. We may make generalizations such as that the presence of another adult during the interview causes negative response effects if the respondent is reporting about threatening personal information, but there will be counter-examples reported so that we deal with probabilities less than one.

There are two chief differences between our analysis and that of a typical survey. For a typical survey, one could have the complete range of information on independent variables for all respondents. That is not the case for the studies we have reported. Most of these studies have been concerned only with the relation between one or a few independent variables and response effects. Sometimes it is possible by careful reading to determine the characteristics of other variables that are not analyzed in the study. For most studies, however, there are a large number of independent variables for which no information is available. Thus, our large sample size of items and studies is misleading, since much of the information is missing. Most of the generalizations made are based on samples much smaller than the total. This, unfortunately, limits our ability to discuss interaction effects in great detail. As the reader will observe, some of these interaction effects appear to be of great importance. At least, these results point to the gaps in information about response effects that may be fruitfully explored in future research.

The other difference between a typical survey and this meta-study is that on a survey, respondents are considered to be equally reliable. In this study, a cursory reading of a few articles is enough to persuade anyone that there are large differences in the quality of the research. We have attempted to quantify the quality of the research by considering the following factors:

a. Researcher's reputation

b. Type of sample

c. Methodological details given in report

d. Type of validating information

Some of the analyses weight the results by this quality measure. In general, however, we have not found that these weighted results differ from the unweighted ones. We also coded the journal in which the report was found, the year of the study, the researcher's professional background, and the size of sample. Our analyses to date have not indicated that these are important variables in explaining differences in response results.

Independent Variables

Effects of Question, Questionnaire Design and Interviewing Situation

Length of interview
Location of interview, presence of others
Subject of report
Threat
Saliency
Method of administration
Structured or unstructured questions
Position of question in questionnaire
Position of question relative to related questions
Deliberate bias in questionnaire wording or deception in experiment
Number of words in question
Difficulty of words in question
Social desirability of answer

Time and Memory Factors

Time period
Records available
Aided recall

Characteristics of Respondents

Age
Sex
Occupation
Income
Education
Race
Religion
Political preference
Household size
Anxiety
Hostility
Conformity
Personal effectiveness
Yeasayer
Mobility

Characteristics of Interviewers

- Age
- Sex
- Education
- Occupation of household head in interviewer's household
- Race
- Religion
- Social class
- Experience
- Training
- Expectations
- Hostility
- Anxiety

Type of Data