

20157



116962

IMPROVING SURVEY DESIGN:
FOCUS GROUP INTERVIEWING,
RANDOMIZED RESPONSE

Laurie Ekstrand

Institute for Program Evaluation
General Accounting Office
Washington, D.C.

Paper prepared for the annual meeting of the Southern Political
Science Association, November 5-7, 1981, Memphis, Tennessee.

019551

Surveys are scrutinized regularly in relation to sampling plans, non-response bias, and analysis issues. Less frequently are concerns raised about the degree to which respondents understand the survey items and the extent to which they respond truthfully. Yet in many cases these are highly significant issues in assessing the validity of the survey results.

Pretesting of the survey instrument is often the means used to provide assurance that the questionnaire works. And this includes assurance that the respondents understand the questions and are able to recount activities or provide opinions that fit appropriate response categories. It is quite possible, however, that respondents can complete a pretest successfully without having a real understanding of the topic(s) of interest. This can be the case especially when respondents pick their answers from categories provided in the survey instrument.

A sizeable body of literature has developed during the last 20 years related to questionnaire design.¹ Attention has been focused on matching questions to the understanding levels of respondents, eliciting truthful responses to sensitive questions, and improving question format and wording. Other related research has assessed the impact of interview mode (telephone vs in-person vs mail) on response patterns and the development of special interviewing procedures for unique respondent groups (i.e. the elderly, children, etc.).

This paper does not attempt to review all techniques that have been developed to improve survey instrument design. It discusses two procedures that have applicability to political science research situations, but are little used in the discipline. Focus group interviewing and randomized response are the two techniques presented. The former is a means for developing the information needed to match respondent understanding level to a survey instrument, and to increase the information known about the survey topic prior to questionnaire design. The latter is one of several methods currently in use to increase respondents' willingness to answer sensitive questions truthfully. Neither of these methods are new, nor are they related to each other. Focus group interviewing is an established part of marketing research (although usually used as an end in itself rather than in conjunction with a survey), and randomized response was first introduced in 1965.²

FOCUS GROUP INTERVIEWING

It is common sense to talk to some likely respondents in the process of designing a survey instrument. Often this is done at the pretest stage of instrument development and sometimes on an informal basis at earlier stages. Very infrequently is it done on a more formal basis as an integral part of the preliminary instrument design process. Yet in many situations it is not reasonable to assume that respondents will adequately understand the survey questions. If

technical terms are involved, if topics are complex, if knowledge of current events is required, or if the respondent population is such that their comprehension level is expected to be below what would normally be expected, it is particularly important to build procedures into the instrument design process that will result in the collection of meaningful data. A focus group interview, or series of interviews, can result in the kinds of information needed to develop a workable data collection instrument.

A focus group, or focus panel, interview is a quasi-structured interview with a group of respondents. Selection of participants (usually from 8 to 12 people) for a session is purposive in nature generally striving for a range of participants typifying the variety of anticipated survey respondents. For example, if the researcher wants to know how persons with diverse political attitudes feel about certain policy issues, respondents would be screened to insure that the group of participants differ along the liberal-conservative spectrum.

The interview session is led by a moderator who guides the discussion through a pre-set series of topic areas. The moderator controls the session to ensure that everyone contributes, that issues of interest are probed and explored sufficiently, and that the discussion does not stray to topics entirely irrelevant to the purpose of the session. The outward appearance of a good group session is that of an informal, unstructured group discussion with no apparent leader. But, in fact, the discussion is under the

direction of the moderator with a specific agenda of items to be covered and information to be obtained.

A moderators guide is prepared for each session or set of sessions. The guide usually takes the form of a topic outline of the subjects to be covered, an indication of the depth in which they should be covered, and sometimes a time limit per section. Often specific probes are included to get needed information if it doesn't develop spontaneously during the interview. Each session starts with a brief introduction to the discussion by the moderator, then short personal introductions from each member of the group to "break the ice." As in a questionnaire, initial topics are intended to make respondents feel comfortable in the interview situation. A focus group usually lasts an hour and a half to two hours. Respondents often lead the discussion from topic to topic in an order entirely different from what was planned in the topic guide. This is no problem as long as all issues are discussed.

Although it is preferable for only one person to moderate each session, one or more unobstrusive observers can attend a session as non-participants. If the interview is not taped (video or audio) it is important for observers to attend the session to take notes on significant points. The moderator would take notes as the session proceeds also although the demands of running the interview often preclude taking extensive notes. Audio taping of sessions is the general practice.

The results of the sessions can be analysed in several ways.

The most thorough and by far most expensive method is to have tapes of the sessions transcribed and then content analysed. Short of transcription, which can be both time consuming and expensive, one, or preferably several analysts can listen to the tapes, pulling out information and organizing their findings around the subjects in the moderator's guide, adding new topics that may come up during the session.

A study of factors affecting tax compliance sponsored by the Internal Revenue Service utilized focus group interviews as part of the pilot survey instrument design process.³ Following are some examples of the kinds of information that were found in the focus group sessions that were helpful in designing the survey:

- Respondents tended to answer questions in relation to taxing policy rather than tax-paying procedures. The latter was the topic of interest. As a result of this, extra emphasis was placed on the correct topic in the introduction to the questionnaire.

- Focus group participants used the terms "exemptions" and "deductions" interchangeably yet they have distinct meanings to IRS. There was a wide range of definitions of an IRS "audit." To some respondents they had been audited if they had received a computerized notice of a math error on their return. Others felt that only an actual examination of a taxpayer's records is an audit. Since the effect of an audit on future compliance was an important issue, it was vital that respondents defined "audit" uniformly. The focus groups exposed numerous seemingly commonplace terms that required precise definitions in the questionnaire.

- Respondents seemed to feel very differently in relation to compliance about paying taxes on their "regular" income (i.e. salary for a full-time job from which taxes were withheld) than on "extra" income (i.e. wages from part-time work, hobbies, interest income, etc.). A series of questions were added to probe this distinction. Neither the IRS nor the research staff had proposed this line of questioning prior to the focus groups.
- Group participants were less reticent about discussing their own non-compliant behavior than had originally been anticipated. For this reason it was decided that the compliance questions for the pilot survey would be asked in a straightforward manner.⁴

Focus group interviews do not provide quantifiable data that can be considered representative of the population as a whole. How, then, can it be advocated that a questionnaire be designed based on information gathered through focus group interviews? First, it is advocated the focus group analysis information be added to knowledge that the researcher already has (i.e. knowledge of other related studies, understanding of questionnaire design techniques, substantive knowledge of the issue area, etc.). In many cases the focus group information will reinforce questionnaire design decisions previously made on the basis of other information.

In other cases the focus group information may suggest different question wordings or entirely different lines of questions than had been envisioned originally. In these situations the potential negative and positive impacts of using the guidance from the focus groups must be weighed against the impacts of not using the

information. Focus group interviews may suggest, perhaps, that some technical terms be more specifically defined than had been originally intended. The potential negative impact of not doing this - that respondents may not understand the questions seems to far outweigh the negative impact of doing it - slightly increasing the length of the questionnaire. In other worrds, the dangers involved in making a Type II error are limited, so it is reasonable to take the risk. In other situations, such as when the focus groups suggest the addition of new lines of questioning that would force dropping other sets of question, common sense has to guide the decision making process.

RANDOMIZED RESPONSE

Randomized response is a method of eliciting truthful responses to sensitive questions in an interview situation. It works by providing the respondent complete confidentiality. Even the interviewer does not know what the respondent answers (or would answer) to the sensitive question. The method is appropriate for situations in which either "yes" - "no" or a minimal number of categorical responses are required, and where the objective is to get an aggregate level estimate of incidence of a sensitive behavior. In other words, it is appropriately used, for example, to ask respondents if they robbed a bank, or how many

times they robbed a bank (within certain categories), but not how they did it.

The technique requires that the sensitive question of interest be paired with a non-sensitive question with a known aggregate response pattern. A means is required also for the respondent to be directed to answer either the sensitive question or the paired non-sensitive alternative without the interviewer knowing which question is being answered. The probabilities of the respondent getting each question (sensitive or non-sensitive) must be known also.

An example can demonstrate the procedure most easily. The questions that follow can be paired to assess the extent to which a group of respondents cheat on their taxes by failing to report all of their income. The decision of which question the respondent answers is left to the outcome of the respondents' toss of a coin (it could be based on the toss of dice, selection of colored balls from a box, etc.). If the toss results in "heads" the respondent answers question A; "tails", question B.

Example Questions

- A. Did you fail to report any of your taxable income on your 1980 tax return?
- B. Did the toss of the coin produce "heads?"

During the interview the procedure is explained to the respondent who is encouraged to examine the coin (or dice, etc.) for reassurance that it is not weighted or otherwise biased. Prior to answering the questions of interest one or more dry-run questions are administered to make sure the respondent understands what to do and is convinced that it is impossible for the interviewer to know which question in each pair is being answered. The actual interview then proceeds with the interviewer recording "yes" or "no" responses to the paired sets of questions.

The analysis of data involves subtracting the number of "yes" responses that were expected from the non-sensitive question and adjusting for the number of respondents who would have answered the non-sensitive question. Some hypothetical response data for the sample set of questions will demonstrate the procedure:

Results: 100 respondents
60 "yes" responses
40 "no" responses

Probability of answering each question: 50%
50 answered coin toss question
50 answered compliance question

Number of "yes" responses attributable to answering the coin toss question: 25
($\frac{1}{2}$ of the 50 respondents who answered the question)

Number of "yes" responses attributable to the compliance question: 35
(total "yes" responses (60) minus those attributable to the coin toss question (25))

Percent of respondents who are noncompliant in reporting taxable income: 70%
(35 of the 50 who responded to the compliance question)

The formula for this computation is shown below:

$$p_s = \frac{p_t - (1-P)p_n}{P}$$

where

p_s = the proportion of respondents who have failed to report all of their income on their 1980 tax return.

p_t = the proportion of total respondents answering "yes" to either of the set of questions.

p_n = the probability of obtaining a "yes" response to the non-sensitive (coin toss) question.

P = The probability of responding to the sensitive (compliance) question based on the randomizing device (the coin toss).

For this example the formula would yield:

$$p_s = \frac{.6 - (1-.5)(.5)}{.5}$$

Randomized response is not without drawbacks. One of the more obvious is that individual respondent characteristics cannot be linked with the same individual's answers to the sensitive questions. It is possible, however, with large samples to analyse data from sub-groups of respondents separately and therefore link some characteristics to behavior. For example, with sufficient sample sizes, data in relation to tax non-compliance could be analysed to determine whether males or females are more non-compliant. The data would be divided into two groups according to respondent's sex and each groups results analysed separately. The outcomes of

the two analyses would then be compared.

A second drawback, which it shares with all sensitive question techniques, is that the relationship between the incidence of admitted behavior and actual behavior is usually unknown. There is evidence that for at least some types of questions randomized response is superior to other methods in eliciting truthful responses. Bradburn and Sudman report findings from a study designed to measure veracity of responses to questions across different modes of inquiry.⁵ The four methods tested were: 1) face-to-face direct questions 2) telephone direct questions 3) self-administered questions and 4) randomized response. The questions included whether the respondent 1) was registered to vote 2) voted in a primary election 3) had a library card 4) had declared bankruptcy, and 5) had been arrested for drunken driving. The first three of these questions are of the type that a respondent would be inclined to answer "yes" when in fact the answer is "no." For the remaining two questions the reverse would be the case. Respondents were selected from public records of bankruptcy, voter registration, etc. so that the correct responses were known to the researchers. Their results suggest that randomized response is effective in increasing the reporting of socially undesirable behaviors but is not effective compared to other methods in reducing overreporting of socially desirable acts.

Another IRS study completed at about the same time as the pilot

compliance survey (cited previously) allows additional comparisons of the effectiveness of randomized response in relation to other methods of inquiry. Both the pilot survey and this second study, the CSR study, asked respondents if they had ever overstated deductions or underreported income on their Federal tax return.⁶ As previously indicated the pilot survey made these inquiries as direct questions. The CSR study used randomized response with a portion of respondents and "locked box" with others.⁷

The CSR randomized response results indicated that 26% of the taxpayer population had underreported income, and 14% overstated deductions. Comparable figures from the pilot survey direct questions (during in-person interviews) were 14% and 7% respectively. Estimates based on IRS audits indicate that the randomized response results approach actual non-compliance rates.⁸

A third potential criticism of randomized response is that the procedure is too complicated - that it might work with sophisticated populations but with less sophisticated respondents it would not work. The evidence, however, seems to weigh against this being the case. The most impressive of this evidence is a report of the successful use of the technique in a study of self-induced abortions among rural Ethiopian women.⁹ Post survey interviews of the respondents indicated a high level of understanding of the procedure, and confidence in the anonymity of their responses.

CONCLUDING COMMENTS

Neither randomized response nor focus group interviewing should be universally adopted by political science survey researchers. In the case of randomized response the applications may be limited. Randomized response seems most useful for assessing incidence of anti-social behaviors or feelings. Political science survey researchers are more frequently confronted with problems related to the over-reporting of socially desirable behaviors such as voting. There are situations, however, in which political scientists may find randomized response useful. For example, political scientists may want to measure the extent of willful non-compliance of city or state administrators with a Federal regulation for which enforcement activity is known to be minimal.

The applications for focus group interviewing may be more widespread. Almost all questionnaire design projects would benefit from discussion of the topics of interest with respondents prior to instrument development. The question here is when the need for focus groups overcomes time and cost constraints.

FOOTNOTES

1. See, for example, Sudman and Bradburn's (1974) book and the same authors (names in reverse order) more recent book (1979). Additionally, the Public Opinion Quarterly regularly publishes articles discussing empirical research related to the improvement of questionnaire design and administration.
2. Stanley L. Warner, "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias." Journal of the American Statistical Association, vol. 60 (1965), pp.63-69.
3. This study was conducted by Westat, Inc. under contract number TIR-78-50. It was designed to develop methodologies to study factors affecting compliance. The focus group interviews and the pilot survey were two of several phases of the project. The author of this paper was the Senior Analyst for Westat on this study.
4. The decision to use a straightforward questioning technique was also predicated on the fact that another IRS sponsored survey proceeding at about the same time would be utilizing randomized response and another sensitive question technique.
5. Norman M. Bradburn and Seymour Sudman, Improving Interview Method and Questionnaire Design (San Francisco: Jossey - Bass Publishers, 1979).
6. The cited study was conducted by CSR Inc. for the Internal Revenue Service in 1979. It consisted of a survey of a national sample of 5000 taxpayers to tap attitudes and behaviors related to a variety of issues associated with paying income taxes.
7. The "locked box" technique involves asking respondents to write their responses to sensitive questions on a piece of paper and to place the paper in a locked box. The box is usually made of lucite or other clear material so the respondent can see that other answer sheets are already in the box. The locked box proved to be the least effective method of eliciting truthful responses to compliance questions.
8. Compliance estimates were based on IRS's 1973 Taxpayer Compliance Measurement Program findings. This work involved analysis of complete audits of a random sample of taxpayer returns.

9. L. P. Chow and others, "Feasibility of the Randomized Response Technique in Rural Ethiopia." American Journal of Public Health, vol. 69 (1979), 273-6.

REFERENCES

- Bradburn, N.M. and Sudman, S. Improving Interview Method and Questionnaire Design. San Francisco: Jossey - Bass, 1979.
- Chow, L.P. and others. "Feasibility of the Randomized Response Technique in Rural Ethiopia.", American Journal of Public Health, 1979, 69, 273-76.
- Sudman, S. and Bradburn, N.M. Response Effects in Surveys: A Review and Synthesis. Chicago: Aldine, 1974.
- Warner, S.L. "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias." Journal of the American Statistical Association, 1965, 60, 63-69.