

handling nonresponse issues

Larry E. Miller
Keith L. Smith

Evaluation studies often use mail questionnaires to obtain evidence. Using information only from those that choose to respond can introduce error. Data gathered from self-selected respondents may not represent the opinions of the entire sample or population. At issue is what to do about nonrespondents. Numerous surveys have been conducted to evaluate programs and their results have been stamped invalid because of this problem.

Evaluation studies often use mail questionnaires to obtain evidence. Using information only from those that choose to respond can introduce error. Data gathered from self-selected respondents may not represent the opinions of the entire sample or population. At issue is what to do about nonrespondents. . . .

Nonresponse Problem

Imagine this situation: An agent is evaluating a workshop on "Farm Financial Management" conducted last year. Although some information was collected at the meetings, a year-end follow-up is believed appropriate since accounts are now being closed out for the year.

Assume that an appropriate evaluation method is used. A delayed posttest would substantiate the earlier data and the overall value of the workshop. Therefore, a survey is to be conducted of the 453 participants. A complete list (frame) of all participants (the population) is made, a random sample of 200 participants is drawn and a valid, reliable, and appropriate questionnaire is mailed to them. By the deadline, 80 replies (40%) are received.

Larry E. Miller: Professor, Department of Agricultural Education, The Ohio State University—Columbus and *Keith L. Smith*: Assistant Professor, Department of Agricultural Education, The Ohio State University—Columbus. Accepted for publication: May, 1983.

Are the data from these 80 truly representative of the 200 to whom the questionnaire was mailed? Are the 80 representative of the 453 participants? Did only those that liked the workshops reply? Did only those who disliked the workshops, and were upset, reply? At this point, the agent doesn't know the answers to these questions. Any or all could be true! If the results were now reported, neither would the reader (supervisor, county commissioner, or legislator) know the true worth of the program.

Unfortunately, this is the point where a lot of people stop. They're stymied! So, they go ahead and complete their report. Some even camouflage this error by neither stating the return rate nor even reporting the number of questionnaires sent out and the number returned. Such evaluators have a nonresponse problem. Their results are only true for those that chose to respond, and may be biased.

The evaluator would have to preempt all statements in the report by saying these conclusions are only true for the "respondents." The charge to the evaluator was to evaluate the effectiveness of the workshop and not just to gather data from those who cared enough to respond. In short, the overall evaluation of the workshop is now invalid.

A lot of time, effort, and dollars have produced results of little value to the agent, the Extension Service, the people served by Extension, or the knowledge base about such programs. How, then, can the nonresponse problem be resolved?

Controlling the Error

The first control strategy involves getting back as many responses as possible. This step can be done by following some tested guidelines on construction of the cover letter and questionnaire.¹

Next, carefully plan a follow-up procedure to encourage response. Sending postcards to announce the questionnaire and as follow-ups to the questionnaire have been successful at improving response rates. A second mailing of the complete questionnaire packet can be made. These procedures have produced high (70%-90%) return rates. Such techniques are particularly effective when the people in the sample know the evaluator. The questionnaire can be mailed many times, with even six mailings showing rewards.²

Additional Strategies

While following these suggested procedures, certain strategies have improved the response rate. Effective techniques that have been employed are using stamped outgoing and return envelopes, using Extension's official letterhead and staff co-signatures, personally signing all letters,

mailing so the questionnaire arrives at a less busy time for the respondents, assuring confidentiality, offering a summary of results, using rewards (gum, mints, pens, stickers) or promised rewards, and specifying in the cover letter a deadline date to receive a response.

Other techniques to consider using are colored paper, mailing all materials flat, keeping questionnaires short, using humorous or poetically novel encouragement, or personal appeals and appeals based on the social benefits of the results.³

Handling the Reminder

Even the most carefully conceived and executed plan isn't likely to produce 100% return. The evaluator still has to provide evidence that the results are true for the sample. Even though a 90% response would provide much greater confidence in the results than 60%, the work isn't done. Here are several strategies the Extension evaluator can use.

Ignore Nonrespondents

A preponderance of surveys seem to employ this approach and result in the evaluation being open to criticism by any or all of its consumers. Obviously, if this strategy were used, a knowledgeable writer would only be able to generalize to the respondents. The findings aren't necessarily true of the sample and, thus, of the population from which it was drawn. This strategy, limiting the generalizability of the results, isn't recommended.

Compare Respondents to Population

The Extension evaluator often has access to information about those served, such as size and type of farming operation, age, sex, socioeconomic status. These characteristics of the population (for example, participants in the Farm Financial Management program) could be compared to those of the respondents. If the respondents are typical of the population (statistical tests can be done), this similarity can be reported and the evaluator can then generalize from respondents to the sample. If there are differences, results must be limited to the respondents. The characteristics used for such comparisons obviously should be related to the variable(s) studied.

Given the Farm Financial Management example, data on the size of the farming operation could be accumulated on all the participants and summarized. Similar data on the respondents would be gathered. If the data for the respondents were similar to those of the population, the assumption could be made that the respondents are a subpopulation of the total population. The rationale would be that the respondents are truly representative of the population. This

rationale is the same that probabilistic (random) sampling provides. Other characteristics would be compared in the same manner.

Compare Respondents to Nonrespondents

Characteristics of the respondents (farm size, farm enterprises, age) can be compared to those of the nonrespondents. If the nonrespondents don't appear different (statistical analysis can be done), then the results can be generalized to the sample and population. If the groups appear different, generalizations would have to be confined to the respondents. Again, appropriate characteristics must be chosen.

To illustrate, if the nature and size of the production enterprises were known, then those data would be summarized for respondents and nonrespondents and the two compared. When combined with other characteristics that are compared in a similar manner, these characteristics would form a basis for trying to generalize. If the groups were similar, the respondents are assumed representative of the sample since they're not different than the nonrespondents and combined both groups comprise the probabilistic sample.

Compare Early to Late Respondents

Research has shown that late respondents are often similar to nonrespondents.⁴ Thus, one way to estimate the nature of the replies of nonrespondents is through late respondents. Late respondents are statistically compared to early respondents using the evaluation data to justify generalizing from the respondents to the sample.

If data on the characteristics are unavailable, available evaluation data can be used with this technique. Respondents can be dichotomized into those that respond early and those that respond late. These two groups can be compared statistically to determine differences between the groups. With late respondents assumed typical of nonrespondents, if no differences are found, then respondents are generalized to the sample. If differences are present, data are weighted proportionately for determining the statistics to describe the sample.

"Double-Dip" Nonrespondents

Once the deadline for submission of the questionnaire is past, a random sample (10%-20%) of the nonrespondents is drawn. Telephone or personal interviews are then used to obtain evaluation data from the "double-dipped" sample using the questionnaire as an interview schedule. These data from the interviews are then statistically compared with the data from the respondents. If the data are similar, the data can be pooled and generalized to the sample/population.⁵

If differences are present, data are weighted proportionately for determining the statistics to describe the sample.

Discussion

If the procedures appear to require more statistical manipulation than would be preferred, help could be sought from a university evaluation, research, or statistical specialist. In reality, the comparisons are quite simple and microcomputers with statistical software or minicomputers can handle these comparisons easily.

The important thing is that one can now eliminate the nonresponse error. This elimination will enable the evaluator to say that the results are true for the sample. If the sample was randomly chosen from an accurate frame, then the results are true for the population. Then, if the data were gathered with appropriate questionnaires, the survey will produce valid results.

The method of drawing the sample with a simple random, stratified, systematic, proportionate, or multistage strategy has to do with sampling error. Nonresponse can be a problem with any sample in a survey.

Double-dipping would produce the most empirically sound procedure, and would be preferred over the other techniques. Comparing early to late respondents on known characteristics, or comparing respondents to the population each necessitate assumptions that leave the results open to question.

Ignoring the nonrespondents is exactly the procedure that has many people questioning the overall validity of survey research. The nonrespondents can't be ignored if the evaluation studies are to be valid.

Steele pointed out that sound procedures must be selected for important judgments concerning programs, and that an easy procedure for evaluation isn't always possible.⁶ Survey techniques that employ mail questionnaires require sound procedures to produce valid results. Decisions about the continuance of a program and personnel matters such as promotion, tenure, and salary are often based on results from such studies. Given the importance of these decisions, the validity of the results is of utmost importance.

Footnotes

1. Donald A. Dillman, *Mail and Telephone Surveys: The Total Design Method* (New York: John Wiley and Sons, 1978) and Seymour Sudman and Norman M. Bradburn, *Asking Questions* (San Francisco: Jossey-Bass, 1982).
2. Frederick Mosteller, "Errors: Nonsampling Errors," in *International Encyclopedia of Social Sciences*, David L. Sills, ed. (Crowell: Collier and Macmillan, Inc., 1968), pp. 118-19.

3. Dillman, *Mail and Telephone Surveys*; Arnold S. Linsky, "Stimulating Responses to Mailed Questionnaires: A Review," *Public Opinion Quarterly*, XXXIX (Spring, 1975), 82-101; Mosteller, "Errors"; David J. Pucel, Howard F. Nelson, and David N. Wheeler, *Questionnaire Follow-Up Returns as a Function of Incentives and Responder Characteristics* (Minneapolis: University of Minnesota, Department of Industrial Education, 1970); and Sudman and Bradburn, *Asking Questions*.
4. John A. Clausen and Robert N. Ford, "Controlling Bias in Mail Questionnaires," *Journal of the American Statistical Association*, XLII (December, 1947), 499; John C. Flanagan and others, *The American High School Student* (Pittsburgh: Project Talent Office, 1964); Herbert Goldhor, "The Use of Late Respondents to Estimate the Nature of Non-Respondents" (Washington, D.C.: ERIC document EDO83309, TM003274, 1974); and Sheldon W. Newman, "Differences Between Early and Late Respondents to a Survey," *Journal of Advertising Research*, II (June, 1962), 37-39.
5. Flanagan, *American High School Student*; Marjorie N. Donald, "Implications of Nonresponse for the Interpretation of Mail Questionnaire Data," *Public Opinion Quarterly*, XXIV (Spring, 1960), 99-114; Eugene C. Hagbert, "Validity of Questionnaire Data: Reported and Observed Attendance in an Adult Education Program," *Public Opinion Quarterly*, XXV (Fall, 1968), 453-56; and Palmer O. Johnson, "Development of the Sample Survey as a Scientific Methodology," *Journal of Experimental Education*, XXVII (March, 1959), 167-76.
6. Sara Steele, "An Emerging Concept of Program Evaluation," *Journal of Extension*, XIII (March/April, 1975), 7-22.