Produced by
Section on Survey Research Methods
American Statistical Association
1429 Duke Street
Alexandria, VA 22314-3415 USA

ASA Series
What Is a Survey?

How to Conduct Pretesting

Section on Survey Research Methods
American Statistical Association

This pamphlet, **How to Conduct Pretesting,** is the fifth in the ASA series ***What Is a Survey?*** It looks at how to check out or pretest a questionnaire—among the most important planning steps in any survey. Care at this point will not guarantee success, but a lack of care will almost always lead to extra costs and a lost opportunity to collect the required information.

The ***What is a Survey?*** series is written primarily for the general public. Its overall goal is to improve survey literacy among individuals who participate in surveys or use survey results. The series is designed to promote a better understanding of what is involved in carrying out sample surveys—especially those aspects that have to be taken into account in evaluating the results of surveys.

How to
Conduct Pretesting

The systematic checking or pretesting of a questionnaire is central to planning a good survey. As mentioned earlier in this series, the survey sponsors should play a major role in developing the data-collection instruments being proposed—including any testing being done. Much of the accuracy and interpretability of the survey results hinge on this pretesting step—**which should never be omitted.**

Pretesting is critical for identifying questionnaire problems. These can occur for both respondents and interviewers regarding question content, "skip patterns," or formatting. Problems with question content include confusion with the overall meaning of the question, as well as misinterpretation of individual terms or concepts. Problems with how to skip or navigate from question to question may result in missing data and frustration for both interviewers and respondents. Questionnaire formatting concerns are particularly relevant to self-

Pretesting is critical for identifying problems for both respondents and interviewers.

administered questionnaires, and if unaddressed, may lead to loss of vital information.

Pretesting is a broad term that incorporates many different methods or combinations of methods.

This pamphlet briefly describes **eight suggested techniques** that can be used to pretest questionnaires. These techniques have different strengths and weaknesses. They can be invaluable for identifying problems with draft questionnaires and also for evaluating surveys in the field.

Types of Pretesting

Pretesting techniques are divided into two major categories—**pre-field** and **field**. Pre-field techniques are generally used during the preliminary stages of questionnaire development. They include respondent focus groups and cognitive laboratory interviews.

Six field techniques that test questionnaires under operational conditions are also covered. These include behavior coding of interviewer/respondent interactions, interviewer debriefings, respondent debriefings, split-sample tests, and the analysis of item nonresponse rates and response distributions.

1. Respondent Focus Groups

Focus groups—a form of in-depth group interviewing—are conducted early in the questionnaire development cycle and can be used in a variety of ways to assess the question-answering process.

Such groups may gather information about a topic before questionnaire construction begins (*for example, to learn how people structure their thoughts about a topic, their understanding of general concepts or specific terminology, or their opinions about the sensitivity or difficulty of the questions*).

Focus groups help identify variations in language, terminology, or interpretation of questions and response options. Self-administered questionnaires can be pretested in a focus group, to learn about the appearance and formatting of the questionnaire. In addition, knowledge of content problems is gained.

One of the main advantages of focus groups is the opportunity to observe a great deal of interaction on a topic in a limited period of time.

They also produce information and insights that may be less accessible without the give and take found in a group. Because of their interactive nature, however, focus groups do not permit a good test of the ''normal'' interviewing process. Researchers also do not have as much control over the process as with other pretesting methods. (F*or example, one or two people in the group may dominate the discussion and restrict input from other focus group members.*)

2. Cognitive Laboratory Interviews

Cognitive laboratory interviews are also generally used early in the questionnaire

development cycle. They consist of one-on-one interviews using a structured questionnaire in which respondents describe their thoughts while answering the survey questions. **"Think aloud" interviews,** as this technique is called, can be conducted either concurrently or retrospectively (*i.e. the respondents' verbalizations of their thought processes can occur either during or after the completion of the questionnaire*).

"Think aloud" cognitive interviews can be conducted either concurrently or retrospectively.

Laboratory interviews provide an important means of finding out directly from respondents what their problems are with the questionnaire. In addition, small numbers of interviews (as few as 15) can yield information about major problems—such as respondents repeatedly identifying the same questions and concepts as sources of confusion. Because sample sizes are not large, repeated pretesting of an instrument is often possible.

After one round of lab interviews is completed, researchers can diagnose problems, revise question wording to resolve these problems, and conduct additional interviews to see if the new questions are better.

Cognitive interviews can incorporate follow-up questions by the interviewer—in addition to respondents' statements of their thoughts. Different types of follow-up questions are used. **Probing questions** are used when the researcher wants to focus the respondent on particular aspects of the question-response task. (*For example, the interviewer may ask how respondents chose their answers, how they interpreted reference periods, or what they thought a particular term meant.*) **Paraphrasing** (i.e., asking the respondents to repeat the question in their own words) permits the researcher to learn whether the respondent understands the question and interprets it in the manner intended. It may also reveal better wordings for questions.

3. Behavior Coding

Behavior coding of respondent-interviewer interactions involves systematic coding of the interaction between interviewers and respondents from live or taped interviews.

The emphasis is on specific aspects of how the interviewer asked the question and how the respondent reacted. When used for questionnaire assessment, the coding highlights interviewer or respondent behaviors indicative of a problem with the question, the response categories, or the respondent's ability to form an adequate response. *For example, if a respondent asks for clarification after hearing the question, it is likely that some aspect of the question caused confusion.* Likewise, if a respondent interrupts before the interviewer finishes reading the question, then the respondent may miss information that might be important to giving a correct answer.

In contrast to pre-field techniques, behavior

coding requires a sample size sufficient to address analytic requirements. *For example, if the questionnaire contains many skip patterns, it is necessary to select a large enough sample to permit observation of various movements through the questionnaire.*

Behavior coding allows systematic detection of questionnaire problems.

The determining sample sizes for behavior coding should take into account the relevant population groups for which separate analyses are desired.

The value of behavior coding is that it allows *systematic detection* of questions that have large numbers of behaviors that reflect problems. It is not usually designed to provide answers about the source of the problems. It also may not distinguish which of several similar versions of a question is better.

4. Respondent Debriefings

Respondent debriefings involve incorporating structured follow-up questions at the end of a field test interview to elicit quantitative and qualitative information about respondents' interpretations of survey questions. For pretesting purposes, the primary objective is to determine whether concepts and questions are understood by respondents in the same way that the survey sponsors intended.

Respondent debriefings can also be used to evaluate other aspects of respondents' tasks, such as their use of records to answer survey questions or their understanding of the purpose of the interview. In addition, respondent debriefings can be useful in determining the reason for respondent misunderstandings. Sometimes results of respondent debriefings show a question is superfluous and can be

eliminated. Alternatively, additional questions may need to be included in the final questionnaire. Finally, the debriefings may show that concepts or questions cause confusion or misunderstanding as far as the intended meaning is concerned. Some survey goals may need to be greatly modified or even dropped.

A critical aspect of a successful respondent debriefing is that question designers and researchers must

Respondent debriefings can be useful in determining the reason for respondent misunderstandings.

have a clear idea of potential problems so that good debriefing questions can be developed. Ideas about potential problems can come from pre-field techniques conducted prior to the field test, from analysis of data from a previous survey, from careful review of questionnaires, or from observation of actual interviews.

Respondent debriefings have the potential to supplement information obtained from behavior coding. As previously discussed,

behavior coding can demonstrate the existence of problems but does not always indicate the source of the problem. When designed properly, the results of respondent debriefings can provide information about the problem sources and may reveal problems not evident from the response behavior.

5. Interviewer Debriefings

Interviewer debriefings traditionally have been the primary method to evaluate field tests. The interviewers who conduct the survey field tests are queried to use their direct contact with respondents to enrich the questionnaire designer's understanding of questionnaire problems.

Interviewer debriefings have traditionally been the primary method to evaluate field tests.

Although important, interviewer debriefings are not adequate as the sole evaluation method. Interviewers may not always be accurate reporters of certain types of questionnaire problems for several reasons:

■ When interviewers report a problem it is not known whether it was troublesome for one respondent or for many.

■ Interviewer reports of problem questions may reflect their own preference for a question rather than respondent confusion.

■ Experienced interviewers sometimes change the wording of problem questions as a matter of course to make them work and may not even realize they have done so.

Interviewer debriefings can be conducted in several different ways:

■ *Group-setting debriefings* are the most common

method, involving a focus group with the field test interviewers.

■ *Rating forms* obtain more quantitative information by asking interviewers to rate each question in the pretest questionnaire on selected characteristics of interest to the researchers (whether the interviewer had trouble reading the question as written and whether the respondent understood the words or ideas in the question, among others).

■ *Standardized interviewer debriefing questionnaires* collect information about the interviewers' perceptions of the problem, prevalence of a problem, reasons for the problem, and proposed solutions to a problem. They can also be used to ask about the magnitude of specific types of problems and to test an interviewer's knowledge of subject-matter concepts.

6. Split-Panel Tests

Split-panel tests refer to controlled experimental testing among questionnaire variants or interviewing modes to determine which is "better" or to measure differences between them. For pretesting multiple versions of a questionnaire there needs to be a previously determined standard by which to judge the differences.

Multiple versions of a draft questionnaire can be tested under controlled experimental conditions in a split–panel test.

Split-panel tests are also used to calibrate the effect of changing questions—particularly important in the redesign and testing of surveys where the comparability of the data collected over time is an issue.

Split-panel tests can incorporate changes in a single question, a set of questions, or an entire questionnaire. It is important to provide for adequate sample sizes in a split-panel test so that differences of substantive interest can be measured well. It is also imperative that these tests involve the use of randomized assignment so differences can be attributed to the question or questionnaire, and not to something else.

7. Analysis of Item Nonresponse Rates

Analysis of item nonresponse rates from the data collected during a field test (involving one or multiple panels) can provide useful information about how well the questionnaire works. This can be done by looking at how often items are missing (item nonresponse rates).

Both response and nonresponse rates provide useful information about how well a questionnaire works.

These rates can be informative in two ways:

- ◼ "Don't know" rates can determine how difficult a task is for respondents to do.
- ◼ Refusal rates can determine how often respondents find certain questions or versions of a question too sensitive to be answered.

8. Analysis of Response Distributions

Analysis of response distributions for an item can be used to determine whether different question wordings or question sequences produce different response patterns. This kind of analysis is most useful when pretesting more than one version of a questionnaire or a single questionnaire in which some known distribution of characteristics exists for comparative purposes.

When looking at response distributions in split-panel tests, the results do not necessarily reveal whether one version of a question produces a better understanding of what is being asked than another. Knowledge of differences in response patterns alone is not sufficient to decide which question best conveys the concept of interest.

At times response distribution analysis demonstrates that revised question wording has no effect on estimates. Response distribution analyses should not be used alone to evaluate modifications in question wording or sequencing. It is useful only in conjunction with other question evaluation methods—such as respondent debriefings, interviewer debriefings, and behavior coding.

Combining Methods

Both pre-field and field testing should be done when time and funds permit; but, there are some situations in which it is not feasible to use all methods. Still, it is particularly desirable to meld the objective with the subjective methods—the respondent centered with the interviewer-centered. This complementarity allows for both good problem identification and problem resolution and provides an evaluation of broad scope.

Combining pre-field and field methods provides an evaluation of broad scope.

Where Can I Get More Information?

In addition to the pamphlets in this series, ASA also makes other brochures available upon request:

■ *Ethical Guidelines for Statistical Practice*

■ *Surveys and Privacy,* produced by the ASA Committee on Privacy and Confidentiality.

For the above brochures or other pamphlets in the *What Is a Survey?* series, contact:

Section on Survey Research Methods
American Statistical Association
1429 Duke Street
Alexandria, VA 22314-3415 USA
(703) 684-1221/fax: (703) 684-2037
Email: asainfo@amstat.org
Web site: http://www.amstat.org/
sections/srms/

Besides the ASA, there are many other associations that are concerned with the proper collection and use of survey data:

■ **The American Association for Public Opinion Research** (AAPOR) offers a number of publications—perhaps the most relevant of these is the one entitled *Best Practices for Survey* and *Public Opinion Research Survey Practices AAPOR Condemns.* To obtain copies, call (313) 764-1555 or visit their Web site at http://www.aapor.org.

■ **The National Council on Public Polls** publishes another useful pamphlet, *Twenty Questions a Journalist Should Ask About Poll Results.* To obtain a copy, call (800) 239-0909.

■ **The Research Industry Coalition, Inc.**, publishes a brochure, *Integrity and Good Practice in Marketing and Opinion Research.* To obtain a copy, call (516) 928-6803.

■ **The Council of American Survey Research Organizations** publishes a pamphlet, *Surveys and You.* To obtain a copy, call (516) 928-6954, or email their Web site at http://www.casro.org.