DOCUMENT RESUME

ED 325 513 TM 015 751

AUTHOR Boser, Judith A.; Clark, Sheldon B.
TITLE Improving the Quality of Questionnaires: A Tool for Evaluators.
PUB DATE Oct 90
NOTE 27p.; Paper presented at the Annual Meeting of the American Evaluation Association (Washington, DC, October 18-20, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Authors; *Check Lists; Evaluation Methods; *Evaluators; Literature Reviews; *Mail Surveys; *Questionnaires; Research Design; *Researchers; *Research Methodology; Test Construction; Test Reliability
IDENTIFIERS *Check List Desirable Characteristics Mail Surveys; Experts

ABSTRACT

A checklist of characteristics of effective mail questionnaires was developed that reflects some degree of consensus among experts in survey research and that can serve as a guide to novices designing a questionnaire. Journal and book sources of information about mail questionnaires were reviewed to yield 83 items in 7 categories that comprised the study survey. The survey was mailed to six authors of books on survey research and six experienced practitioners of survey research. Only one author failed to reply; the response rate was 92%. Detailed background information was provided by 10 of the 11 individuals who participated in this phase of the study. Eight of the 83 items were judged desirable for all mail questionnaires, possibly because of a lack of clarity in the instrument. A revised instrument was submitted to a validation panel of 10 individuals experienced in survey research as well as to 8 of the original 11 participants. Of the 64 items that a majority of the participants in the first phase of the study would usually recommend, 38 were supported by the validation panel, with 80% or higher indicating that they also would usually make the recommendation. The attached "Check List of Desirable Characteristics of Mail Surveys" is a compilation of the items usually recommended by at least 87.5% of the first group of experts and 80% of the validation panel. These results indicate that questionnaire design may be a science up to a certain point, but beyond that point it is an art. A bibliography of the 21 journal and book sources that were reviewed is provided. The survey instrument used in this study is included. (SLD)

Improving the Quality of Questionnaires: A Tool for Evaluators

by

Judith A. Boser
Bureau of Educational Research and Service
The University of Tennessee
212 Claxton Education Building
Knoxville, TN 37996-3400

Sheldon B Clark
Science/Engineering Education Division
Oak Ridge Associated Universities
P.O. Box 117
Oak Ridge, TN 37831-0117

Paper Presented at Evaluation '90, the Annual Meeting of
the American Evaluation Association
Washington, DC
October 18-20, 1990

IMPROVING THE QUALITY OF QUESTIONNAIRES: A TOOL FOR EVALUATORS

Background and Objective

Mail surveys are used frequently, particularly in higher education institutions (Fuqua, Hartman, & Brown, 1982). According to Babbie (1973), "survey research is probably the best known and most widely used research method in the social sciences today. . . . To some extent, everyone in the United States at least has been affected by surveys" (p. i). While the research instrument is only one component of the overall research endeavor, in mail surveys the questionnaire takes on added importance. The potential respondent encounters it in isolation, with no interviewer present to encourage the respondent to participate or to provide explanations. The individual must be motivated to complete the questionnaire, and the questionnaire must be designed to facilitate the respondent's providing valid responses. "The task required of respondents must appear to be easy and attractive. . . . Anything [respondents] particularly dislike about the layout, wording, or emphasis of the questions may deter them (Hoinville, 1978, p. 127)." Berdie, Anderson, and Niebuhr (1986) concur, noting that "poorly constructed formats [(the physical arrangement of questions on the page)] influence not only response rates but also the quality of responses obtained" (p. 23).

Not surprisingly, one of the most commonly used techniques for collecting evaluation data is the paper-and-pencil questionnaire. Because the value of the data--and the evaluation itself--is at least partially dependent on the quality of the data collection device, it is important that evaluators be knowledgeable about principles of questionnaire design. As Riecken (1972, p. 94), points out, "the development of measuring devices is a technical problem of social science rather than one peculiar to evlauation studies." The literature on social science survey research is a particularly relevant resource when evaluation data are collected by mail.

The body of literature on survey research includes a number of works that describe various authors' systems of total survey design--including the design of the instrument itself, sampling considerations, choosing questions, cover letters, follow-up procedures, etc. Also in the literature is a plethora of studies dealing with various effects (e.g., on response rate, on turnaround time) of alternate forms of one or more elements of questionnaire design (e.g., multiple choice versus open-ended responses, variations in type size or type style). One would be hard pressed, however, to identify a simple check list of characteristics of effective mail questionnaires that might provide useful guidance for the evaluator. The total system concept is too restrictive for this purpose: not only do many of the suggestions offered represent a single point of view, but the guidelines may be so specific that they are difficult to generalize to a situation other than that for which they are illustrated. The problem with recommendations found in most journal articles is that they are likely

to be too narrow in scope for the purpose described. The objective of the present study was to develop a check list of characteristics of effective mail questionnaires (a) that reflects some degree of consensus of experts in survey research, and (b) that can be used as a general guideline by novice questionnaire designers.

Phase One

Instrument Development

Significant journal and book sources that provide general guidelines for designing mail questionnaires were identified through an investigation of current books in print and ERIC listings. Those sources containing guidelines only for other survey techniques (e.g., telephone or face-to-face interviews) were not included unless it could be determined that the recommended procedures were equally applicable for mail surveys. The final list of sources is attached as the Bibliography.

The specific recommendations from each of these books or articles were listed and categorized. Only those characteristics which were deemed desirable by several of the authors were retained; those characteristics mentioned by only a few authors and those about which there was disagreement were excluded from the list. The remaining list of desirable characteristics was edited to exclude redundancies and re-categorized independently by the two authors. The authors then compared and discussed their respective lists, ultimately producing the 83 items and seven categories contained in the Phase One instrument. Throughout their discussions the authors recognized that there was more than one way in which items could be categorized, and that the categorization of items and the labels selected were somewhat arbitrary.

The final grouping of items resulted in seven categories requiring varying numbers of responses. The categories and their respective numbers of items (responses) were as follows: General Appearance (14 items); Instructions (8 items); Choice of Items (8 items); Order of Items (15 items); Item Format (16 items); Choice of Response Options (10 items); and Wording (12 items).

Once the characteristics had been selected, it was apparent that some might be more important than others. To this end, the respondents were asked to indicate, for each item, the extent to which the characteristic would be recommended for mail survey questionnaires using the following ratings: "all" ("recommended for all mail survey questionnaires"); "some" ("recommended for some but not all mail surveys"); or "none" ("not recommended"). There was space following each section under a heading of "other" for the respondents to add other characteristics that they thought should have been included in that section.

The questionnaire was photocopied and assembled in booklet format (7" by 8-1/2") using two sheets of ivory colored, legal-sized paper that were printed on both sides, folded, collated, and saddle-stitched. The front page served as a cover and the last (or eighth) page was reserved for

comments. Identical directions for responding to the items were placed at the top of each of the six inside pages.

Participants

It was considered important that the participants in this study be knowledgeable and experienced in survey research and represent various research environments. Six authors of books on survey research (from the attached bibliography) were invited to participate. One declined. Their publications contained guidelines for the total development of mail questionnaires (including wording, order, and format or layout) and were not specific to a particular research emphasis (academia, public opinion polls, marketing research). Six experienced practitioners of survey research were selected from the membership of the American Educational Research Association's (AERA's) Special Interest Group on Survey Research in Education on the basis of their activities in the group and their survey research background.

Detailed background information was provided by ten of the eleven individuals who participated in this phase of the study. In addition to the five who were authors of books on survey methodology, each of the remaining six had made formal presentations on issues of survey research methodology at national professional conferences. Each of the five authors is in a leadership position in an organization which has a focus on survey research. Four of the other participants are employed in postsecondary institutions in units that focus on research and/or evaluation.

The ten who supplied background information had amassed a total of 172 years of experience in survey research, with individual experience varying from 7 to 40 years (median = 15.5 years). All ten had carried out surveys during the previous year, and most considered it a typical year. These researchers had conducted from 1 to 30 surveys themselves, as well as providing consultation on others. The research focus and the target population varied both within and across individuals. The major types of surveys were described as public opinion, needs assessment, program evaluation/effectiveness, and institutional, consumer, and attitude studies. Target groups enumerated included the following: the general public; program participants; students; alumni; consumers; client groups; various occupational groups, including professionals (e.g., judges, lawyers); and groups of employees within organizations (e.g., supervisors, managers).

Procedures

A copy of the instrument, an explanatory cover letter, and a pre-stamped reply envelope were mailed in the spring of 1988. One follow-up reminder which included another copy of the instrument was mailed to each of the nonrespondents approximately one month after the initial mailing. Only one of the twelve potential participants in this phase of the study did not respond, for a response rate of 92 percent.

Results

Only eight of the 83 items were judged to be desirable for "all" mail questionnaires by all of the respondents. Some items lacked clarity, as indicated by the comments of the respondents. There were two items that none of the respondents recommended for "all" mail questionnaires. Comments from the researchers' professional colleagues led them to conclude that the poles of the initial response categories ("all" and "none") may have been too extreme, causing participants to reject them for the only other available option, "some" ("recommended for some but not all mail surveys") if they thought of one or more applications for which the recommendation would or would not be made. Limitations of the instrument were thus considered as a possible factor contributing to the lack of consensus among the experts. Phase Two of the study was then undertaken.

Phase Two

Instrument Revision

Concern about effects of the response options led the researchers to revise the instrument. The response option "all" ("recommended for all mail survey questionnaires") was broadened to "usually=usually or always recommended for mail survey questionnaires" while the other extreme, originally stated as "none" ("not recommended"), was revised to become "seldom=seldom or never recommended for mail survey questionnaires." The middle category of "some" ("recommended for some but not all mail surveys") was changed only to reflect the tone and phrasing of the new response options, becoming "sometimes=sometimes recommended for mail survey questionnaires."

Almost all of the items from Phase One were used in the revised questionnaire. Four of the five items which had caused confusion to participants in the previous study were rewritten in an effort to clarify them, and the fifth such item was deleted.

The revised instrument contained 82 items in seven categories: General Appearance (14 items), Instructions (8 items), Choice of Items (7 items), Choice of Response Options (10 items), Wording (12 items), Order of Items (15 items), and Item Format (16 items).

Two versions of the questionnaire were produced. The items themselves were identical, but in one version, following each section of items (a section contained items of one of the seven types), participants were asked to indicate the circumstances or types of surveys in which items rated as "sometimes" would be recommended. If there was not sufficient blank space at the bottom of the page, the facing page was left blank for this purpose. In the other version, there was no space for participants to explain or list circumstances relevant to items they had rated as "sometimes." A copy of the second-version questionnaire is appended to this paper.

Questionnaires of the first type were duplicated on blue legal size paper that was stapled in the middle to form a booklet. Because of the additional space required for explanatory comments, the

questionnaire had 12 pages. Questionnaires of the second type were green and required only eight pages.

Participants

The researchers wanted to take advantage of the extraordinary credentials of the 11 experts surveyed in Phase One, but they were concerned about the potential bias of the results caused by that group's prior exposure to the instrument. For this reason, a second group, a "validation panel," was identified.

A group size of 10 was deemed desirable for this validation panel. Oversampling was initiated by sending questionnaires to 15 individuals to obtain 10 participants for the validation panel. The respondents in the validation group were the first 10 individuals (from the 15 who were sent questionnaires) who returned completed questionnaires. In all, a total of 12 from the validation sample returned the survey instruments without benefit of follow-ups. Background information was provided by nine of the 10 validation panel members.

For the validation group, experience in survey research activities varied from five to 30 years, with a median of 13 years. Participants had been involved in conducting from none (n=1) to eight (n=1) mail surveys in 1988, with a mean of three. Five of the nine participants indicated 1988 was typical of their survey activities, but the remaining four indicated they were usually more active in survey research. Eight of the nine had conducted and reported studies of survey or questionnaire methodology, and four had published articles or books on survey methodology. Eight of the 10 individuals were involved in institutional or organizational research, and five of them limited their activities to this type. Other types of survey research included public opinion (n=2), consumer (n=1) and social science (n=1). Populations surveyed included members of organizations (n=5), program participants (n=5), alumni (n=4), and the general public (n=3). Seven of the 10 were college/university faculty members, with three employed by research institutes within their institutions. Two individuals were employed by research divisions or sections in large organizations, and one individual was in a public school research division.

In addition to the validation panel, the 11 participants from Phase One were given an opportunity for continued involvement in the project. Of these 11, eight completed questionnaires. Two others indicated that they thought their responses would be biased by their participation in Phase One and declined to participate. One of the two recommended a colleague to participate in his place, and the other forwarded his questionnaire to a colleague. Both recommended individuals became part of the validation group. Only one of the original 11 participants failed to respond in any manner.

Procedures

The 12-page survey forms, cover letters and postage-paid return envelopes (along with previously promised copies of the results of Phase One) were mailed to the original 11 participants

in the spring of 1989. The cover letter emphasized the change in response options. One follow-up mailing, containing a letter, a replacement copy of the questionnaire and a stamped, addressed return envelope, was sent approximately one month later, Completed survey instruments were received from eight of the 11 original participants (73 percent).

Eight-page survey forms, cover letters and postage-paid return envelopes were mailed to the 15 individuals selected for the validation sample in the summer of 1989. Responses from 12 of them constituted an 80 percent response rate.

Analysis

Frequency distributions were prepared for all items for the Phase One participants and for the validation panel. Explanatory comments listing special circumstances in which practices would sometimes be recommended were noted for items on which there was not total agreement by participants from Phase One.

Items were listed in four groups, based on responses of the Phase One participants: items on which all who responded marked the item as "usually" recommended, items on which all but one who responded marked the item as "usually" recommended, items on which all but two who responded marked the item as "usually" recommended, and the items that lacked general acceptance as usual practices.

After items were grouped according to responses by the original participants, the percentage of validation panel members who would "usually" recommend each item was calculated. If an item was "usually" recommended by 80 percent or more of the validation panel, the item was considered to have been supported.

Phase Two Results and Discussion

On 26 of the 82 items (32 percent), all eight of the original participants agreed that they would "usually" recommend the item for mail survey questionnaires. And on another 8 items, all of those responding to the item (n=6 or n=7) rated the items as usually recommended. Those items are listed below.

On 25 of the 34 items, the validation panel agreed at the 80 percent level or higher. Those items appear in bold-face type in the list that follows. The 9 items that do not appear in bold-face type were supported by fewer than 80 percent of the validation panel.

Items Recommended by All of Original Group

A. GENERAL APPEARANCE

1. The title of the study/questionnaire is likely to appeal to the survey population.
2. Instrument looks easy to complete.
6. Type is clear and legible.
11. Appreciation for completing the instrument is expressed.

9a. The front page (or cover) contains the study/instrument title, prominently displayed.

B. INSTRUCTIONS

2. Instructions are brief.
3a. Instructions are clear: They specify when to put a check mark and when to write in a response.
3b. Instructions are clear: They indicate whether multiple responses are allowed.

6. If items appear on both sides of the page, an indication is given that the instrument continues on the reverse side (e.g., "please turn over").

C. CHOICE OF ITEMS

2. Each item seeks just one piece of information.
3. All items are essential and relevant to the purposes of the survey.
4b. For items used for skip/filter/screen purposes, instructions are few and simple.

D. CHOICE OF RESPONSE OPTIONS

1a. Response options exhaust all possibilities or include "other," "undecided," or "neutral" category.
1d. Response options do not contain more than one alternative that could be correct unless multiple responses are allowed.
1g. Response options are appropriate for the item.

1f. Response options are brief.
2b. Items with Likert-type response options use a balanced scale. (n=7)

E. WORDING

1. The choice of words is appropriate to the literacy level of the survey population.
3d. Items are simple, direct, and unambiguous. They <u>do not</u> contain instances of double negatives in items and/or response options. (n=7)
3e. Items are simple, direct, and unambiguous. They <u>do not</u> contain instances of negatively worded items coupled with agree/disagree response format. (n=7)

Items Recommended by All of Original Group (continued)

F. ORDER OF ITEMS

1b. The initial items are applicable to all members of the survey population.
1d. The initial items are nonthreatening.
1e. The initial items are interesting.
5. If reference is made to a previous item, that item appears on the same page or on the facing page.
6. Items with similar content are grouped together; within each content group, items with the same response format are presented together.

1a. The initial items are clearly connected to the stated purpose of the survey.
4. Classification or demographic information is solicited at the end of the instrument unless needed for screening purposes.
7b. Within a topic/content area, the items progress from most familiar to least familiar.
7c. Within a topic/content area, the items progress from least objectionable to most objectionable.

G. ITEM FORMAT

8. Response options are arranged vertically (or in columns if several consecutive items use the same response options). (n=6)
9. Response options are close to the item stem. (n=7)
11. There is adequate space for responding. (n=7)
13. When ranking, the number of items to be ranked is limited (e.g., three best and three worst). (n=5)

10. The space for responding to items is on the same side of the page throughout the instrument. (n=6)

On another 18 items, only one of the original (Phase One) participants rated the item as "sometimes" or "seldom," while the rest accorded it the "usually" rating. Comments or special circumstances from those original participants are shown in italics to the right of the item.

Using the same 80 percent agreement, 12 of these items were supported by the validation sample and are again shown in bold-face type.

Items Usually Recommended by All But One of Original Group

Item	Circumstances/Comments	
A3.	**Margins are adequate; instrument doesn't look crowded.**	*Crowding preferable to longer instrument.*
A5.	**Printing does not bleed through the paper.**	
A8.	There are not too many variations in size and style of type.	*Interesting, varied format may add appeal for children, teens, or specialized audiences.*
B5.	**The tone of the directions is polite (e.g., "please").**	
C1.	**The respondent is able to provide answers to the questions in the instrument.**	*Person should have the knowledge; sometimes opinions of naive persons sought (although they don't feel able to provide the answers).*
C4c.	For items used for skip/filter/screen purposes, instructions appear immediately after the response options. (n=6)	
D1b.	**Response options are mutually exclusive.**	*Except items with"check all that apply."*
E3b.	**Items are simple, direct, and unambiguous. They do not contain instances of "loaded" items (that use emotionally colored words). (n=6)**	*Such words are appropriate in attitude surveys.*
E3c.	Items are simple, direct, and unambiguous. They do not contain instances of assumption of an existing state of affairs (e.g., "Do you still...").	NOTE: Wording of this item was confusing to some participants.
E3g.	**Items are simple, direct, and unambiguous. They do not contain instances of "giveaway" words (e.g., "all").**	NOTE: Wording of this item was confusing to some participants.
F1c.	**The initial items are easy. (n=6)**	*More important that initial items be interesting if there are no items that are both interesting and easy.*
F7d.	**Within a topic/content area, the items progress from objective to subjective.**	*No clear agreement on this.*

Items Usually Recommended by All But One of Original Group (continued)

Item		Circumstances/Comments
F8.	**Items that require recall are organized by logical time sequence. (n=6)**	*No clear agreement on this.*
G2.	If necessary, either sublettering (e.g., 4a, 4b, 4c) or numbering by sections (i.e., starting each section with item 1) is used to limit the apparent number of items.	
G3.	**Each item and its response options are on the same page.**	*Long items (25 Likert items) may not fit on a page.*
G4.	Statements or questions, rather than phrases, are used in collecting demographic information (e.g., "How old were you on your last birthday?" instead of "Age."). (n=6)	*Level of detail needed and literacy of reader must be considered.*
G14b.	**For check lists, column headings are carried over from one page to another. (n=6)**	*Better not to have to carry over to second page.*
G14c.	For check lists, column headings are presented parallel, rather than perpendicular, to the item stem. (n=5)	*There may be too many options at times.*

For the following 12 items, all but two of the Phase One participants agreed that they should be recommended. Only one of the items, G7, was validated at the 80 percent level.

Items Usually Recommended by All But Two of Original Group

Item	Circumstances/Comments	
A7.	Size and style of type used for headings is consistent throughout the instrument. Consistency is also evident for items and response options.	
A9b.	The front page (or cover) contains general directions.	*Preferably in cover letter.*
B4.	Instructions are visually different from the body of the instrument (e.g., in size and/or style of type).	
D1e.	Response options include both sides of issue or question.	*Both or neither; some issues may have more than two viewpoints*
E2.	Both sides of issue (or neither side) are included in the item stem.	*There may be more than two sides. Some items may focus on one side of an issue.*
E3f.	Items are simple, direct and unambiguous. They do not contain instances of qualifying clauses, especially at end of stem.	NOTE. Wording of this item was confusing to some participants.
E3i.	Items are simple, direct, and unambiguous. They do not contain instances of vague terminology (e.g., "the country," "just," "fair," "you")	*Many of the terms acceptable for response categories. This statement doesn't specify what part of the item the words are to be omitted from.*
F7a.	Within a topic/content area, the items progress from general to specific.	*Sometimes may not care if earlier responses influence summary item. No clear agreement.*
G1.	Items are numbered with Arabic numerals. (n =5)	
G5.	If an item stem requires two or more lines, the second and subsequent lines are indented. (n=5)	
G7.	**When response options are provided (including, if appropriate, a response option of "other"), each response option has either a numeric or alphabetic code beside it.** (n=5)	*Recommended to facilitate computer data entry. Otherwise checking beside or circling response may be acceptable.*
G12.	Open-ended items are used sparingly. (n=5)	*If most likely responses are known.*

The following items appear to be more controversial or highly situational in nature than commonly accepted, or the relevant items in this study were not clearly written. These items would not be included among a generic list of questionnaire characteristics. Some of the circumstances identified by the respondents are listed.

Non-Generic, or Situation-Specific Items

Item		Circumstances/Comments
A4.	Paper is white or light-colored with dark ink.	*Brighter colors, varied formats may be more appealing to children, teens, specific audiences.*
A9c.	The front page (or cover) contains the name of the sponsor.	*Some prefer this in cover letter. Response rate may be lower if very personal/confidential information sought.*
A9d.	The front page (or cover) contains the address of the sponsor.	*Some prefer this in cover letter or at end of questionnaire. Complete address not needed in organization survey.*
A10.	For a multi-page questionnaire, the back page does not contain items but may be used for comments.	*Use for questions if needed rather than adding pages. Less important if comments have been sought throughout the questionnaire.*
B1.	General instructions that apply to the entire instrument are provided at the beginning of the instrument.	
B3c.	Instructions are clear. They provide guidance for expected length of open-ended responses.	*If there is a need to limit length. Available space is an indicator.*
C4a.	For items used for skip/filter/screen purposes, the use of this type is justified.	*One justification is to shorten the overall questionnaire. Avoid when possible in mail surveys.*
C4d.	For items used for skip/filter/screen purposes items pertaining to only some of the respondents are indented beneath the filter question.	*This format may be less appropriate for map-like or flowchart-based formats.*
D1c.	Response options include a "don't know" option.	*Some prefer to force respondents to make a choice.*

Non-Generic, or Situation-Specific Items (continued)

Item		Circumstances/Comments
D2a.	Items with Likert-type response options have an appropriately labeled midpoint.	*Some prefer not to include a midpoint or to label only the end points.*
D3.	Sensitive information (e.g., age, salary) is collected using ranges for response options	*Unless interval level data needed. Ranges may be preferable if anonymity or confidentiality is a concern.*
E3a.	Items are simple, direct, and unambiguous. They do not contain instances of jargon, technical terms, or uncommon abbreviations.	*Technical terms can be used if those in the sample would be familiar with them (engineering terms for a survey of engineers).*
E3h.	Items are simple, direct, and unambiguous. They do not contain instances of inexact words or phrases (e.g., "any," "most," "several," "usually," "often," "regularly," "much the same").	*These are acceptable response categories, and this item doesn't state whether these words are to be omitted from the stem or response.*
E3j.	Items are simple, direct, and unambiguous. They do not contain instances of the word "questionnaire" or "check list" in heading or text.	
F2.	If there are any sensitive or difficult items, they appear in the middle or near the end of the instrument, but not at the very end.	*All items may be sensitive. Place at end unless critical to study (more commitment to answer because of time already spent on the questionnaire).*
F3.	Open-ended items appear last.	*Should follow items if used to clarify or expand responses to them.*
G6.	The respondent is asked to circle or underline responses.	*Except when listing responses might influence respondents or when possible responses cannot be predicted.*
G14a.	For check lists, if long, a line is skipped after every three to six items.	*May not be needed if items go across most of the page or if there is sufficient space between items.*

There was much less agreement on the use of items in section G than in other sections. There was also more reluctance to rate the items in that section, possibly indicating confusion over the items themselves.

Summary

There were 34 items on which all (or all who rated the item) of the Phase One experts agreed that they would usually recommend. For 25 of those items, there was 80 percent or higher support from the validation panel that the items should be included in a list of recommendations usually made in mail surveys. Of the 18 items on which all but one of the Phase One panel supported, 12 of the validation panel provided support. And on the 12 items on which all but two of the Phase One panel would usually recommend, only one of the items was supported by the validation panel. In summary, of the 64 items which a majority of the Phase One participants would usually recommend, 38 of them were supported by the validation group at the 80 percent or higher level indicating they also would usually make the recommendation. As consensus declined within the Phase One participants, the support of the validation group also declined.

There are 18 items from the instrument that appear to be recommendations that would be made only in certain circumstances or were poorly written and confusing to participants. In a few cases, participants indicated they would be more likely not to make such a recommendation than to make it.

The method of analysis for these data was arbitrary. The level-of-agreement criterion for the Phase One participants (i.e., all but two or fewer agreeing the recommendation would usually be made) and the 80 percent criterion for the validation panel may be too lenient. The comments from the original sample regarding items that are situation-specific or confusing in themselves may lead to improvement of the wording or intent of some items and the acceptance that there are conditions under which others are applicable.

The attached "Check List of Desirable Characteristics of Mail Questionnaires" is a compilation of those items that were "usually recommended" by at least 87.5 percent of the Phase One experts (7 of 8) and at least 80 percent of the validation panel. Note that the nature of most of these generally-agreed-upon characteristics is very general. Based on this study, it appears that while there are some mail questionnaire recommendations that can usually be made with some degree of confidence, there are other aspects of questionnaire design that are less commonly accepted, and their proper use may depend on the experience and knowledge of the researcher regarding not only questionnaire design but also the population to be surveyed, the information sought, and the circumstances. In other words, questionnaire design may be a science only up to a certain point; beyond that point it is an art, and it would appear that point is reached somewhere prior to the completion of the questionnaire design.

References

Babbie, E.R. (1973). Survey research methods. Belmont, CA: Wadsworth Publishing Co.

Berdie, D.R., Anderson, J.F., & Niebuhr, M.A. (1986). Questionnaires: design and use (2nd ed.). Metuchen, N.J.: Scarecrow Press, Inc.

Boser, J. A., & Clark, S. B. (1990, April). Consensus on desirable characteristics of mail questionnaires: illusion or reality? Paper presented at the annual meeting of the American Educational Research Association, Boston.

Clark, S. B., & Boser, J. A. (1989, March). Seeking consensus on empirical characteristics of effective mail questionnaires: A first step. Paper presented at the annual meeting of the American Educational Research Association, San Francisco. (ERIC Document Reproduction Service No. ED 306 305)

Davis, B.G., & Humphreys, Sheila. (1985). Evaluating intervention programs: Applications from women's programs in math and science. New York: Teachers College Press.

Fink, A., & Kosecoff, J. (1978). An evaluation primer. Beverly Hills: Sage Publications.

Fuqua, D.R., Hartman, B.W., & Brown, D.F. (1982). Survey research in higher education. Research in Higher Education, 17(1), 69-80.

Hoinville, R.J., & Associates. (1978). Survey research practice. London: Heinemann Educational Books.

Riecken, H.W. (1972). Memorandum on program evaluation. In C.H. Weiss (Ed.), Evaluating action programs: Readings in social action and education (pp. 85-104). Boston: Allyn and Bacon.

Suchman, E.A. (1967). Evaluative research. New York: Russell Sage Foundation.

Bibliography

Alwin, D.F. (1982). Survey design and analysis: current issues. Beverly Hills: Sage.

Babbie, E.R. (1973). Survey research methods. Belmont, CA: Wadsworth Publishing Co.

Belson, W.A. (1981). The design and understanding of survey questions. Gower.

Berdie, D.R., Anderson, J.F., & Niebuhr, M.A. (1986). Questionnaires: design and use (2nd ed.). Metuchen, N.J.: Scarecrow Press, Inc.

Bradburn, N.M., Sudman, S., & Associates. (1979). Improving interview method and questionnaire design: response effects to threatening questions in survey research. San Francisco: Jossey-Bass Publishers.

Converse, J.M., & Presser, S. (1986). Survey questions: handcrafting the standardized questionnaire. Beverly Hills: Sage.

Covert, R.W. (1984). A checklist for developing questionnaires. Evaluation News, 5(3), 74-78.

Dillman, D.A. (1978). Mail and telephone surveys: the total design method. New York: John Wiley & Sons.

Erdos, P.L. (1970). Professional mail surveys. New York: McGraw-Hill.

Fink, A., & Kosecoff, J. (1985). How to conduct surveys: a step-by-step guide. Beverly Hills: Sage.

Fowler, F.J. (1984). Survey research methods. Beverly Hills: Sage.

Hoinville, R.J., & Associates. (1978). Survey research practice. London: Heinemann Educational Books.

Labaw, P. (1982). Advanced questionnaire design. Cambridge, MA: Abt Books.

Lockhart, D.E. (Ed.). (1984). Making effective use of mailed questionnaires. New Directions for Program Evaluation (No. 21). San Francisco: Jossey-Bass Publishers. (Especially the following chapters: Improving Mailed Questionnaire Design, by Seymour Sudman & Norman Bradburn; and The Importance of Adhering to Details of the Total Design Method (TDM) for Mail Surveys, by Don A. Dillman, Joye J. Dillman, and Carole J. Makela.)

Moser, C.A., & Kalton, G. (1972). Survey methods in social investigation. New York: Basic Books, Inc.

Nixon, J.E. (1954). The mechanics of questionnaire construction. Journal of Educational Research, 67(7), 481-487.

Oppenheim, A.N. (1966). Questionnaire design and attitude measurement. New York: Basic Books, Inc.

Payne, S.L. (1951.) The art of asking questions. Princeton, NJ: Princeton University Press.

Rossi, P.H., Wright, J.D., & Anderson, A.B. (Eds.). (1983). Handbook of survey research. New York: Academic Press.

Schuman, H., & Presser, S. (1981). Questions and answers in attitude surveys: experiments on question form, wording, and context. New York: Academic Press.

Sudman, S., & Bradburn, N.M. (1982). Asking questions: a practical guide to questionnaire design. San Francisco: Jossey-Bass Publishers.

Checklist of Desirable Characteristics of Mail Questionnaires

GENERAL APPEARANCE

- The title of the study/questionnaire is likely to appeal to the survey population.
- Instrument looks easy to complete.
- Margins are adequate; instrument doesn't look crowded.
- Printing does not bleed through the paper.
- Type is clear and legible.
- Appreciation for completing the instrument is expressed.

INSTRUCTIONS

- Instructions are brief.
- Instructions are clear:
  - They specify when to put a check mark and when to write in a response.
  - They indicate whether multiple responses are allowed.
- The tone of the directions is polite (e.g., "please").

CHOICE OF ITEMS

- The respondent is able to provide answers to the questions in the instrument.
- Each item seeks just one piece of information.
- All items are essential and relevant to the purposes of the survey.
- For items used for skip/filter/screen purposes, instructions are few and simple.

CHOICE OF RESPONSE OPTIONS

- Response options:
  - Exhaust all possibilities or include "other," "undecided," or "neutral" category.
  - Are mutually exclusive.
  - Do not contain more than one alternative that could be correct unless multiple responses are allowed.
  - Are appropriate for the item.

WORDING

- The choice of words is appropriate to the literacy level of the survey population.
- Items are simple, direct, and unambiguous. They do not contain instances of:
  - "Loaded" items (that use emotionally colored words).
  - Double negatives in items and/or response options.
  - Negatively worded items coupled with agree/disagree response format.
  - "Giveaway" words (e.g., "all").

ORDER OF ITEMS

- The initial items are:
  - Applicable to all members of the survey population.
  - Easy.
  - Nonthreatening.
  - Interesting.
- If reference is made to a previous item, that item appears on the same page or on the facing page.
- Items with similar content are grouped together; within each content group, items with the same response format are presented together.
- Within a topic/content area, the items progress from objective to subjective.
- Items that require recall are organized by logical time sequence.

ITEM FORMAT

- Each item and its response options are on the same page.
- Response options are arranged vertically (or in columns if several consecutive items use the same response options).
- Response options are close to the item stem.
- There is adequate space for responding.
- When ranking, the number of items to be ranked is limited (e.g., three best and three worst).
- For checklists, column headings are carried over from one page to another.

Demographic Information

Please provide answers to the following questions regarding your background and experience in survey research activities.

1. How many years have you been involved in survey research activities?
   _____

2. How many mail surveys were you involved in conducting during 1987?
   _____

3. Would you consider 1988 a typical year in regard to your survey activities?
   ____Yes
   ____No, less than usual
   ____No, more than usual

4. Have you ever conducted and reported any studies of survey or questionnaire methodology to find more effective ways of conducting surveys?
   ____Yes
   ____No

5. Have you ever published an article or book on survey methodology or results of your research regarding survey methodology?
   ____Yes
   ____No

6. What type of mail surveys are you generally involved in?
   ____Consumer research
   ____Public opinion polls
   ____Institutional/organization research
   ____Other_____

7. What type(s) of population do you generally survey?
   ____General public
   ____Alumni
   ____Program participants
   ____Users of a particular product
   ____Members of an organization or specific group, such as employees
   ____Other_____

8. What is the nature of your employment?
   ____College/University faculty
   ____Research Institute in a college/university
   ____Employed by large organization as part of research division or section
   ____Survey consultant (private enterprise)
   ____Other_____

Thank you for sharing your experience and expertise with us in this research.

Please return to Judy Boser, The University of Tennessee, 212 Claxton, Knoxville, TN 37996.

DESIRABLE CHARACTERISTICS OF MAIL QUESTIONNAIRES

INSTRUCTIONS: *Listed on the following pages are some generally agreed upon characteristics of effective mail questionnaires. Please indicate the relative importance of each characteristic for mail survey questionnaires by circling your response to the right of the item on the following basis:*

USUALLY *= usually or always recommended for mail survey questionnaires*
SOMETIMES *= sometimes recommended for mail survey questionnaires*
SELDOM *= seldom or never recommended for mail survey questionnaires*

Comments may be added at the bottom of pages on which items appear.

JUDITH A. BOSER

The University of Tennessee

SHELDON B. CLARK

Oak Ridge Associated Universities

Please indicate the relative importance of each characteristic for mail survey questionnaires by circling your response to the right of the item on the following basis:
USUALLY = usually or always recommended
SOMETIMES = sometimes recommended
SELDOM = seldom or never recommended

A. General Appearance

1. The title of the study/questionnaire is likely to appeal to the survey population..... USUALLY SOMETIMES SELDOM
2. Instrument looks easy to complete..... USUALLY SOMETIMES SELDOM
3. Margins are adequate; instrument doesn't look crowded..... USUALLY SOMETIMES SELDOM
4. Paper is white or light-colored with dark ink..... USUALLY SOMETIMES SELDOM
5. Printing does not bleed through paper..... USUALLY SOMETIMES SELDOM
6. Type is clear and legible..... USUALLY SOMETIMES SELDOM
7. Size and style of type used for headings is consistent throughout the instrument. Consistency is also evident for items and response options..... USUALLY SOMETIMES SELDOM
8. There are not too many variations in size and style of type. .... USUALLY SOMETIMES SELDOM
9. The front page (or cover) contains:
   a. the study/instrument title, prominently displayed..... USUALLY SOMETIMES SELDOM
   b. general directions..... USUALLY SOMETIMES SELDOM
   c. the name of the sponsor..... USUALLY SOMETIMES SELDOM
   d. the address of the sponsor..... USUALLY SOMETIMES SELDOM
10. For a multi-page questionnaire, the back page does not contain items but may be used for comments..... USUALLY SOMETIMES SELDOM
11. Appreciation for completing the instrument is expressed..... USUALLY SOMETIMES SELDOM

Please indicate the relative importance of each characteristic for mail survey questionnaires by circling your response to the right of the item on the following basis:
USUALLY = usually or always recommended
SOMETIMES = sometimes recommended
SELDOM = seldom or never recommended

G. Item Format

1. Items are numbered with Arabic numerals..... USUALLY SOMETIMES SELDOM
2. If necessary, either sublettering (e.g., 4a, 4b, 4c) or numbering by sections (i.e., starting each section with item 1) is used to limit the apparent number of items..... USUALLY SOMETIMES SELDOM
3. Each item and its response options are on the same page..... USUALLY SOMETIMES SELDOM
4. Statements or questions, rather than phrases, are used in collecting demographic information (e.g., "How old were you on your last birthday?" instead of "Age")..... USUALLY SOMETIMES SELDOM
5. If an item stem requires two or more lines, the second and subsequent lines are indented..... USUALLY SOMETIMES SELDOM
6. The respondent is asked to circle or underline responses already presented rather than write them on a blank..... USUALLY SOMETIMES SELDOM
7. When response options are provided (including, if appropriate, a response option of "other"), each response option has either a numeric or alphabetic code beside it..... USUALLY SOMETIMES SELDOM
8. Response options are arranged vertically (or in columns if several consecutive items use the same response options)..... USUALLY SOMETIMES SELDOM
9. Response options are close to the item stem..... USUALLY SOMETIMES SELDOM
10. The space for responding to items is on the same side of the page throughout the instrument..... USUALLY SOMETIMES SELDOM
11. There is adequate space for responding..... USUALLY SOMETIMES SELDOM
12. Open-ended items are used sparingly..... USUALLY SOMETIMES SELDOM
13. When ranking, the number of items to be ranked is limited (e.g., three best and three worst)..... USUALLY SOMETIMES SELDOM
14. For checklists:
    a. If long, a line is skipped after every three to six items..... USUALLY SOMETIMES SELDOM
    b. Column headings are carried over from one page to another..... USUALLY SOMETIMES SELDOM
    c. Column headings are presented parallel, rather than perpendicular, to the item stem..... USUALLY SOMETIMES SELDOM

(Please continue to page 8)

Please indicate the relative importance of each characteristic for mail survey questionnaires by circling your response to the right of the item on the following basis:
USUALLY = usually or always recommended
SOMETIMES = sometimes recommended
SELDOM = seldom or never recommended

F. Order of Items

1. The initial items are:
   a. clearly connected to the stated purpose of the survey..... USUALLY SOMETIMES SELDOM
   b. applicable to all members of the survey population..... USUALLY SOMETIMES SELDOM
   c. easy..... USUALLY SOMETIMES SELDOM
   d. nonthreatening..... USUALLY SOMETIMES SELDOM
   e. interesting..... USUALLY SOMETIMES SELDOM
2. If there are any sensitive or difficult items, they appear in the middle or near the end of the instrument, but not at the very end..... USUALLY SOMETIMES SELDOM
3. Open-ended items appear last..... USUALLY SOMETIMES SELDOM
4. Classification or demographic information is solicited at the end of the instrument unless needed for screening purposes. ... USUALLY SOMETIMES SELDOM
5. If reference is made to a previous item, that item appears on the same page or on the facing page..... USUALLY SOMETIMES SELDOM
6. Items with similar content are grouped together; within each content group, items with the same response format are processed together..... USUALLY SOMETIMES SELDOM
7. Within a topic/content area, the items progress from:
   a. general to specific..... USUALLY SOMETIMES SELDOM
   b. most familiar to least familiar..... USUALLY SOMETIMES SELDOM
   c. least objectionable to most objectionable..... USUALLY SOMETIMES SELDOM
   d. objective to subjective..... USUALLY SOMETIMES SELDOM
8. Items that require recall are organized by logical time sequence..... USUALLY SOMETIMES SELDOM

Please indicate the relative importance of each characteristic for mail survey questionnaires by circling your response to the right of the item on the following basis.
USUALLY = usually or always recommended
SOMETIMES = sometimes recommended
SELDOM = seldom or never recommended

B. Instructions

1. General instructions that apply to the entire instrument are provided at the beginning of the instrument..... USUALLY SOMETIMES SELDOM
2. Instructions are brief..... USUALLY SOMETIMES SELDOM
3. Instructions are clear:
   a. They specify when to put a check mark and when to write in a response..... USUALLY SOMETIMES SELDOM
   b. They indicate whether multiple responses are allowed .. USUALLY SOMETIMES SELDOM
   c. They provide guidance for expected length of open-ended responses..... USUALLY SOMETIMES SELDOM
4. Instructions are visually different from the body of the instrument (e.g., in size and/or style of type)..... USUALLY SOMETIMES SELDOM
5. The tone of the directions is polite (e.g., "please"). USUALLY SOMETIMES SELDOM
6. If items appear on both sides of the page, an indication is given that the instrument continues on the reverse side (e.g., "please turn over")... USUALLY SOMETIMES SELDOM

C. Choice of Items

1. The respondent is able to provide answers to the questions in the instrument..... USUALLY SOMETIMES SELDOM
2. Each item seeks just one piece of information. ... USUALLY SOMETIMES SELDOM
3. All items are essential and relevant to the purposes of the survey..... USUALLY SOMETIMES SELDOM
4. For items used for skip/filter/screen purposes
   a. The use of this type is justified..... USUALLY SOMETIMES SELDOM
   b. Instructions are few and simple..... USUALLY SOMETIMES SELDOM
   c. Instructions appear immediately after the response options... USUALLY SOMETIMES SELDOM
   d. Items pertaining to only some of the respondents are indented beneath the filter question..... USUALLY SOMETIMES SELDOM

Please indicate the relative importance of each characteristic for mail survey questionnaires by circling your response to the right of the item on the following basis:
USUALLY = usually or always recommended
SOMETIMES = sometimes recommended
SELDOM = seldom or never recommended

D. Choice of Response Options

1. Response options:
   a. exhaust all possibilities or include "other," "undecided," or "neutral" category..... USUALLY SOMETIMES SELDOM
   b. are mutually exclusive..... USUALLY SOMETIMES SELDOM
   c. include a "don't know" option..... USUALLY SOMETIMES SELDOM
   d. do not contain more than one alternative that could be correct unless multiple responses are allowed..... USUALLY SOMETIMES SELDOM
   e. include both sides of issue in question..... USUALLY SOMETIMES SELDOM
   f. are brief..... USUALLY SOMETIMES SELDOM
   g. are appropriate for the item..... USUALLY SOMETIMES SELDOM
2. Items with Likert-type response options:
   a. have an appropriately labeled midpoint.. USUALLY SOMETIMES SELDOM
   b. use a balanced scale..... USUALLY SOMETIMES SELDOM
3. Sensitive information (e.g., age, salary) is collected using ranges for response options..... USUALLY SOMETIMES SELDOM

Please indicate the relative importance of each characteristic for mail survey questionnaires by circling your response to the right of the item on the following basis:
USUALLY = usually or always recommended
SOMETIMES = sometimes recommended
SELDOM = seldom or never recommended

E. Wording

1 The choice of words is appropriate to the literacy level of the survey population... USUALLY SOMETIMES SELDOM
2 Both sides of an issue (or neither side) are included in the item stem..... USUALLY SOMETIMES SELDOM
3 Items are simple, direct, and unambiguous. They do not contain instances of any of the following pitfalls:
   a. jargon, technical terms, or uncommon abbreviations..... USUALLY SOMETIMES SELDOM
   b. "loaded" items (that use emotionally colored words)... USUALLY SOMETIMES SELDOM
   c assumption of an existing state of affairs (e.g., "Do you still...") ... USUALLY SOMETIMES SELDOM
   d. double negatives in items and/or response options... USUALLY SOMETIMES SELDOM
   e. negatively worded items combined with agree/disagree response format. ... USUALLY SOMETIMES SELDOM
   f. qualifying clauses, especially at end of stem.. ... USUALLY SOMETIMES SELDOM
   g "giveaway" words (e.g., "all"). ... USUALLY SOMETIMES SELDOM
   h inexact words or phrases (e.g., "any," "most," "several," "usually," "often," "regularly," "much the same").. USUALLY SOMETIMES SELDOM
   i. vague terminology (e.g., "the country," "just," "fair," "you") .. ... USUALLY SOMETIMES SELDOM
   j the word "questionnaire" or "checklist" in heading or text..... USUALLY SOMETIMES SELDOM