



How Unclear Terms Affect Survey Data

Floyd Jackson Fowler, Jr.

The Public Opinion Quarterly, Vol. 56, No. 2 (Summer, 1992), 218-231.

Stable URL:

<http://links.jstor.org/sici?sici=0033-362X%28199222%2956%3A2%3C218%3AHUTASD%3E2.0.CO%3B2-D>

The Public Opinion Quarterly is currently published by Oxford University Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/oup.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

HOW UNCLEAR TERMS AFFECT SURVEY DATA

FLOYD JACKSON FOWLER, JR.

Abstract Although writing clear questions is accepted as a general goal in surveys, procedures to ensure that each key term is consistently understood are not routine. Researchers who do not adequately test respondent understanding of questions must assume that ambiguity will not have a large or systematic effect on their results. Seven questions that were drawn from questions used in national health surveys were subjected to special pretest procedures and found to contain one or more poorly defined terms. When the questions were revised to clarify the definition of key terms, significantly different estimates resulted. The implication is that unclear terms are likely to produce biased estimates. The results indicate that evaluation of survey questions to identify key terms that are not consistently understood and defining unclear terms are ways to reduce systematic error in survey measurement.

Introduction

It has been axiomatic at least since Stanley Payne in 1951 wrote his classic book, *The Art of Asking Questions*, that survey questions should be clear. Ideally, they should mean the same thing to all respondents, and they should mean the same thing to respondents as to the researcher, the person who will interpret the answers.

Although the principle seems clear, putting this guideline into practice may be more difficult. The main step that most researchers take

FLOYD J. FOWLER, JR., is senior research fellow at the Center for Survey Research, University of Massachusetts—Boston. Some of the data in this article were presented at the Fifth Conference on Health Survey Research Methods, Keystone, Colo., 1989. This project was supported by grant HS 05616 from the National Center for Health Services Research and Health Care Technology Assessment. This analysis emerged from a research project that, in addition to the author, was conducted by Charles F. Cannell, Graham Kalton, Lois Oksenberg, and Katherine Bischooping at the Survey Research Center, University of Michigan.

to evaluate their questions is the pretest. The most common strategy for pretesting survey instruments is probably to have a few experienced interviewers take 20–40 interviews using a pretest interview schedule and then participate in a debriefing meeting with the investigators. Although there have been protocols outlined for pretesting (Converse and Presser 1986; DeMaio 1983), even at their best, pretests tend to be subjective and unsystematic. The criteria for what interviewers are looking for and the standards they should use for when a “problem” exists are seldom made explicit. Each interviewer has only a small sample of respondents, so the base for evaluating questions is small. Indeed, the pretest puts the interviewer in a complex dual role of interviewer and observer, and it is not clear how well interviewers can play both roles simultaneously. To quote Presser (1989, p. 35): “Although it is likely that pretested questions are better than unpretested questions, there is ever-mounting evidence that many pretested items perform poorly.”

Much of that evidence results from efforts to involve cognitive psychologists in survey research methodology (Jabine et al. 1984; Lessler 1987). When standard survey questions are examined in a laboratory context, ambiguities frequently are uncovered (Royston 1989; Willis, Royston, and Bercini 1989). Debriefing of respondents in a field setting also routinely reveals inconsistent understanding of terms (Belson 1981).

Ambiguous terms pose problems for interviewers in surveys. The rules of standardization limit the extent to which they can or should provide clarification of unclear terms (Fowler and Mangione 1990). The standardized response interviewers are taught that saying, “Whatever it means to you,” may be appropriate for terms that describe subjective states—such as “problems,” “satisfaction,” or “happy”—but this seems inadequate for questions about factual events or experiences such as “visits to doctors” or “having arthritis.”

The most important issue, however, is the way question ambiguity affects the quality of survey data. There are two ways that ambiguous terms can affect the resulting data. If the effect of ambiguity on answers is unsystematic, there is more error in the measurement and the ability to reach valid conclusions about relationships will be reduced, but the estimates of means and the overall distributions may not be affected. If the ambiguity systematically leads respondents to err in one direction, by definition, estimates based on the answers will be biased. Although neither kind of error is desirable, most researchers probably would be particularly concerned about unclear questions that produced biased estimates—estimates that consistently differed from the true value of what they were trying to measure.

This paper emerged from my efforts to develop more systematic ways to use pretests to identify problem survey questions, including those that are not understood consistently by respondents. As will be described below, building on the work of Cannell (Cannell, Fowler, and Marquis 1968; Cannell, Lawson, and Hausser 1975; Cannell and Oksenberg 1988) and Morton-Williams (Morton-Williams 1979; Morton-Williams and Sykes 1984), the behaviors of interviewers and respondents during interviews were coded for clues to question problems. Among the question problems identified were some unclear terms or concepts. In those cases, the questions were revised to improve the definition of key terms and then tested again, using comparable samples and procedures. Hence, that research provided an opportunity to look at the effect of having unclear terms in questions on the estimates that emerge from surveys.

Method

The data for this paper come from a study designed to evaluate strategies for pretesting survey questions. A 60-item survey instrument was constructed, using questions drawn from health survey instruments conducted by government agencies or academic survey organizations. Although a true sampling of such survey items proved impossible, questions were chosen that were judged to be representative of questions commonly used in health surveys of the general public. Items were chosen to cover the range of things typically measured in health surveys, including utilization of medical care, health status, health beliefs, and behaviors related to health.

Samples for all the testing described in this paper were drawn from telephone subscribers in southeastern Michigan. Since the purpose was to evaluate pretesting procedures and not to make general population estimates, strict probability sampling procedures were not used. Interviewers were permitted to substitute a new number after three unsuccessful attempts to contact a household or when someone refused to be interviewed. The protocol for respondent selection within a household called for selecting among those who were at home at the time of contact. However, the procedures for contacting households and substituting and choosing respondents were the same for all phases of the project so that results from the various tests could be combined and compared.

In the initial phase of the study, the test health interview survey was used for 110 pretest interviews. Sixty of these were done by interviewers from the University of Michigan's Survey Research Center and 50,

by interviewers from the Center for Survey Research at the University of Massachusetts—Boston. All interviews were tape-recorded, with respondent permission, of course.

The behaviors during the pretest interviews were then coded by using a scheme adapted from that reported by Cannell and Oksenberg (Cannell, Lawson, and Hausser 1975; Cannell and Oksenberg 1988). The purpose of the coding was to identify questions that appeared to pose problems for interviewers or respondents that were reflected in their behaviors during the interviews. The behaviors that proved to be particularly critical were the rates at which respondents asked interviewers for clarification of questions and the rates at which respondents gave at least one inadequate answer to a question. When either of these behaviors occurred in at least 15 percent of the pretest interviews, the question was considered likely to have a wording problem that deserved attention.

Numerous kinds of problems were identified through this procedure, as is discussed elsewhere (Cannell et al. 1989; Fowler 1989; Oksenberg, Cannell, and Kalton 1991). The focus of this paper is that set of questions that seemed to pose a problem for respondents primarily because a key term or concept in the question was not well defined.

Of the 60 questions, there were seven that seemed to be problematic primarily because of a poorly defined term or concept. These seven questions were revised in an attempt to solve the definitional problem and communicate a consistent understanding of the terms to all respondents. In the revisions, the question objectives were not purposely changed, though in some cases they had to be refined. Table 1 presents the exact wording of the original questions and the revisions. In five of the seven cases, the revision primarily consisted of adding words to the question to clarify some aspect of what was intended. In the case of the question about visits to doctors at an HMO, a single question was turned into two questions. In the case of the question about the number of visits to doctors in the past year, the definition of what does and does not count as a doctor visit proved to be so complicated that a four-question series was substituted for the single question. However, the objectives and what was to be measured remained basically unchanged.

These revised questions became part of a revised survey instrument that was used in a second round of pretesting. In the second phase of pretesting, 150 respondents were interviewed, drawn from the same sample frame as the first phase, using identical respondent selection procedures. Interviews again were split between the staffs of the University of Massachusetts and the University of Michigan, with those staffs taking 50 and 100 interviews, respectively. For the second round,

60 interviews were tape-recorded, and the behaviors during the interview were coded to evaluate the impact of the changes on the behavior of interviewers and respondents.

The focus of the analysis presented below is the extent to which the revised questions, designed to clarify ambiguous terms, also changed the distribution of responses.

Table 1. Original and Revised Questions Designed to Clarify Unclear Terms

Original Question	Problem	Revision
1. During the past 4 weeks, beginning Monday (DATE 4 WEEKS AGO) and ending this past Sunday (DATE 4 WEEKS AGO), have you done any exercise, sports, or physically active hobbies?	What counts as exercise?	In the past 4 weeks, beginning Monday (DATE 4 WEEKS AGO) and ending this past Sunday (LATE LAST SUNDAY), have you played any sports, done any hobbies involving physical activity, or done any exercise, including walking?
2. Do you exercise or play sports regularly?	What counts as exercise?	Do you do any sports or hobbies involving physical activities, or any exercise, including walking, on a regular basis?
3. What is the average number of days each week you have butter?	Does margarine count as butter?	The next question is just about butter. Not including margarine, what is the average number of days each week you have <i>butter</i> ?
4. (Following a question on the number of days per week eggs are eaten): What is the number of servings in a typical day?	How many eggs constitute a serving? What does "a typical day" mean?	On days when you eat eggs, how many eggs do you usually have?

Table 1 (*Continued*)

Original Question	Problem	Revision
5. During the past 12 months, since January 1, 1987, how many times have you seen or talked with a doctor or assistant about your health? Do not count any time you might have seen a doctor while you were a patient in a hospital, but count all other times you actually saw or talked to a medical doctor of any kind about your health.	What counts as a visit to doctor?	Have you been a patient in a hospital overnight in the past 12 months, since July 1, 1987? 5a. (Not counting when you were in a hospital overnight): During the past 12 months, since July 1, 1987, how many times did you actually <i>see</i> any medical doctor about your own health? 5b. During the past 12 months, since July 1, 1987, were there any times when you didn't actually <i>see</i> the doctor, but saw a nurse or other medical assistant working for the doctor? (How many times?) 5c. During the past 12 months, since July 1, 1987, did you get any medical advice, prescriptions, or results of tests over the telephone from a medical doctor, nurse, or medical assistant working for a doctor? (How many times?)

Table 1 (Continued)

Original Question	Problem	Revision
6. (Referring to the place where respondent reported last seeing a doctor): Was that place a health maintenance organization or health care plan (that is, a place you go for all or most medical care, which is paid for by a fixed monthly or annual amount)?	What is an HMO?	Do you belong to an HMO or health plan that has a list of people or places you go to, in order for the plan to cover your health care costs? Was your last visit to a medical doctor covered by your health care plan?
7. During the past 12 months, that is, since January 1, 1987, <i>about</i> how many days did illness or injury keep you in bed more than half of the day? (Include days while an overnight patient in a hospital.)	What counts as "in bed more than half the day"? How specific does the diagnosis have to be to count?	The next question is about extra time you have spent in bed because of illness or injury (including time spent in the hospital). During the past 12 months, since July 1, 1987, on about how many <i>days</i> did you spend several extra hours in bed because you were sick, injured, or just not feeling well?

NOTE.—The original questions were drawn from the core of the National Health Interview Survey and the Health Promotion and Disease Prevention supplement to the NHIS, the Arizona Health Survey conducted for the Center for Health Administration, University of Chicago, by Louis Harris Associates, and a pretest version of the Canadian Fitness Survey conducted by the Canadian Fitness and Lifestyle Institute. In some cases, questions were modified to fit the context or flow of this interview; two questions were adapted from a self-administered questionnaire for use in a telephone survey.

Results

Table 2 presents the distributions of answers to the original and revised questions. For five of the seven questions, answers to the revised question were significantly different from answers to the original one. The other two questions produced differences in the predicted direction that did not quite make the .05 standard for statistical significance.

The fact that the answers to the revised questions were different does not necessarily mean that they were more accurate. However, for at least six of the seven questions that produced changed results, there was a clear *a priori* reason for predicting that the observed change reflected improved measurement. (a) It seems most plausible that the increased report of exercise in two questions stems from the fact that some respondents who walk for exercise excluded it in answers to the original questions. (b) Specifying that margarine does not count as butter explains the decrease in the rate of butter consumption. (c) The changed reporting with respect to egg consumption is certainly more interpretable, as it is clear that some people consider a serving of eggs to be one egg, while others consider it to be two. (d) The increase in the number of visits to doctors reported in response to the revised series is almost certainly more complete than the reporting in response to the single question. (e) Disability days seem likely to be underreported, so that an upward change is likely also to be in the direction of greater accuracy. The only change that is ambiguous with respect to whether or not it constitutes an improvement is obtaining medical care from an HMO. Although the investigators thought that the revised question was clearer, there was not the same basis for estimating a direction of change that would be expected from the clarification that existed for the other questions.

However, the data presented in table 3 provide additional evidence about whether or not the questions were improved. All of the questions that were revised were classified as "problems," either because 15 percent of the original pretest respondents asked for clarification or 15 percent initially gave an inadequate answer to the question. Table 3 compares the rates of these two behaviors in the original and revised interview schedules. It can be seen that, in general, rates of requests for clarification and rates of inadequate answers declined, as one would predict if the questions had been improved. It is particularly worth noting that the HMO questions showed significant improvement on both indicators.

The one question that deteriorated from the perspective of the coding of behavior during the interview pertained to disability days. While the marked increase in reported disability days was predicted, it was

Table 2. Responses to Original and Revised Questions (Percentages)

Question	Yes	No	Total			p^a
Any exercise/4 weeks:						
Original	67	33	100			
Revision	82	18	100			<.05
Exercise regularly:						
Original	48	52	100			
Revision	60	40	100			<.10
	0	1	2-6	7	Total	p
Number days/week have butter:						
Original	33	13	29	23	100	
Revision	55	14	22	9	100	<.01
	1	2	3+		Total	p
Number eggs (servings)/day eaten:						
Original	80	15	5		100	
Revision	33	61	6		100	<.01
	0	1-2	3+		Total	Mean p
Doctor visits/year:						
Original	22	33	45		100	3.6
Revision	19	24	57		100	6.1 <.05
	Yes	No	Total			p
Last doctor visit at HMO:						
Original	39	61	100			
Revision	23	77	100			<.05
	0	1-7	8+		Total	Mean p
Disability days/year:						
Original	57	36	7		100	2.6 <.10
Revision	48	33	19		100	4.0 ^b

NOTE.— N (original) = 110; N (revised) = 150.

^a Based on two-tailed t -tests for differences between means for continuous variables and between proportions for categorical variables. Multicategory variables were dichotomized based on the proportion in the largest category versus all others.

^b Includes one person who reported 90 days, almost twice as many as the next person in either sample. Removing that person reduces the mean to 3.4. In neither case does the difference between means meet the .05 level of significance.

Table 3. Requests for Clarification and Inadequate Answers to Original and Revised Questions (Percentage of Interviews)

Question	Requests for Clarification	Inadequate Answers
Any exercise/4 weeks:		
Original	8	17
Revision	10	13
Exercise regularly:		
Original	5*	20
Revision	0	12
Number days/week have butter:		
Original	18	15
Revision	13	12
Average servings butter/day eaten:		
Original	17**	7
Revision	4	17
Number of eggs (servings)/day eaten:		
Original	33**	12**
Revision	0	0
Doctor visits/year:		
Original	30	17
Revision (4 questions) ^a	24	43**
Last doctor visit at HMO:		
Original	17*	27
Revision (2 questions)	2	18 ^b
Disability days/year:		
Original	15	7
Revision	17	30**

NOTE.—*N* (original) = 110; *N* (revised) = 60.

^a Figures are the sum of behavior on the four questions.

^b 6 and 12 percent, respectively, for the two questions.

* $p < .05$ based on two-tailed *t*-tests for differences in proportions.

** $p < .01$.

not predicted that respondents would give significantly more inadequate answers in response to the revised question.

The series of questions regarding visits to doctors cannot be directly compared to the original question because the revision provided four times as many chances to give inadequate answers and to ask for clarification as the original. In general, the data suggest that the revised, more completely defined questions were also better questions as indicated by requests for clarification and the adequacy of answers.

Discussion

Rasinski (1989) and Schuman and Presser (1981) provide ample documentation that small changes in question wording can affect the distribution of answers. For example, Schuman and Presser found that many more people would "not allow" Communists to speak in the United States than would "forbid" them to do so. However, most of the comparisons that they give pertain to the measurement of subjective phenomena (opinions, knowledge, and feelings) rather than the reporting of factual events. The goal in measuring subjective phenomena is to order people. There are no right or wrong answers, and different distributions in answers to similar questions do not necessarily imply error in either. However, that is not the case when measuring factual phenomena. Although some health services research goals involve the measurement of subjective phenomena, the preponderance of measurement in health research tends to be factual in nature. Researchers want to count events or they want to classify people accurately according to their experiences or behaviors. When a respondent asks an interviewer what is meant by "a serving of eggs" or by "exercise," it is not sufficient for the interviewer to say, "Whatever it means to you." That is a nondirective response, but it does not serve the interests of gathering accurate data. When the researcher interprets the answers, assumptions will be made about the way in which respondents interpreted the questions.

The questions examined here were not the work of amateurs. All were designed by government or academic survey experts, and five of the seven questions presumably had been subjected to standard pretests. It is obvious in retrospect that these questions contain major ambiguities, though presumably it was not obvious to their sponsors and those who pretested them. The fact that in most cases revised questions that clarify and define key concepts produced significantly different distributions of answers is compelling evidence that there was considerable error in the responses to the original questions, and error was almost certainly reduced by the revised wording.

Of course, there is a critical relationship between clear question objectives and clear questions. In fact, error can only be assessed in relationship to a clearly defined measurement goal. One reason for unclear questions, no doubt, is that researchers have not thought through exactly what they want to measure or are unaware of an ambiguity. Using the examples in the test questions, exercise could have been defined to include or exclude walking for exercise; butter consumption could include or exclude margarine. So long as the criteria are not specified in the question, however, some people will answer the question one way, others will answer it another, and there will be

a resulting over- or underestimate of whatever the researcher thinks is being counted or wants to count. One good reason for identifying unclear questions is to alert researchers to places where they need to clarify their objectives.

There is, of course, no basis for saying the revised questions are error free. The question on disability days provides a good example of how a limited revision of a question may not address the total problem posed by a question. Even though this question has been used for years in the National Health Interview Survey and elsewhere, we think it poses major problems of several types. The revised question tried to clarify what being in bed for half a day due to injury or illness meant. In this case, we took some latitude in defining what seemed to be a poorly defined concept; it may still be poorly defined in the revised question, and others might disagree with our sense of what was meant by the question. To the extent that is true, it only reinforces the ambiguity of the question for all who would try to answer it or interpret those answers. Furthermore, we think the task given the respondent is almost impossible: recalling numerous, sporadic, usually minor events over a 1-year period. Clarifying the definition of a disability day should be helpful to understanding but, if anything, the clarification may have made the response task harder. Some question objectives require a completely changed strategy in order to produce questions that can be answered accurately and with ease.

These complexities should not obscure the main implications of the results. First, if one wants to have respondents provide accurate counts of events or behaviors, all respondents should be given adequate definitions of what it is they are supposed to report. Moreover, it is not acceptable to give interviewers definitions to be used only if respondents ask questions or act confused. Consistent measurement requires that all respondents be exposed to the same consistent definitions (Fowler and Mangione 1990).

An obvious concern is that adding detailed definitions to questions will make them long, cumbersome, and perhaps harder for respondents to understand. Sometimes that is a problem, and the benefits of providing definitions must be weighed against the cost. When definitions are very complex, breaking a single question into several questions, as was done with the question on visits to doctors, may be the solution. For those who worry that additional questions will take more interview time, the only response is that researchers always have to weigh the value of each question. However, consider the implicit choice: asking a single question that respondents do not understand and that will, as a consequence, produce biased estimates versus asking three clear questions that will produce more accurate estimates. Certainly, there are many occasions when the three-questions option will be preferable.

Furthermore, there should be no occasions when a researcher does not want to find out that a question is unclear so that the options can be explicitly considered.

To that end, researchers need to do a better job of identifying ambiguous terms and concepts when they are pretesting questions. The old-style pretests are not adequate to the task. The recent emphasis on laboratory testing of questions prior to field-pretesting them is a laudable development that needs to become even more common. Laboratory work to date has shown that interviewing only a few respondents, even fewer than 10, can identify significant ambiguities in question wording. The coding of interviewer and respondent behavior during pretests provides another way to identify questions that include poorly defined terms.

Of course, requests for clarification and inadequate answers can result from question problems other than unclear terms; after coding behavior, there is still a diagnostic task of isolating the aspect of the question that is causing the problem. However, as we have seen, the ambiguities discussed in this paper all showed up in the coding of interview behavior in pretests; the indicators of question problems were, in most cases, reduced by the clarifications. Multiple strategies for identifying unclear terms, involving laboratory tests and field pretests, probably are needed and appropriate.

The most important message from this study is that a commitment to identifying unclear terms in questions and providing adequate definitions is an essential part of good survey research design. When measuring objective phenomena, telling interviewers to say, "Whatever it means to you," is not an adequate substitute for writing clear, well-defined questions.

References

- Belson, William A. 1981. *The Design and Understanding of Survey Questions*. London: Gower.
- Cannell, Charles F., Floyd J. Fowler, and Kent Marquis. 1968. *The Influence of Interviewer and Respondent Psychological and Behavioral Variables on the Reporting in Household Interviews*. Vital and Health Statistics, series 2, no. 26. Washington, D.C.: Government Printing Office.
- Cannell, Charles F., Sally A. Lawson, and Doris Hausser. 1975. *A Technique for Evaluating Interviewer Performance*. Ann Arbor: Survey Research Center, University of Michigan.
- Cannell, Charles F., and Lois Oksenberg. 1988. "Observation of Behaviour in Telephone Interviewers." In *Telephone Survey Methodology*, ed. R. Groves et al., pp. 475-95. New York: Wiley.
- Cannell, Charles F., Lois Oksenberg, Floyd J. Fowler, and Graham Kalton et al. 1989. "New Techniques for Pretesting Survey Questions." Final report to National

- Center for Health Services Research and Health Care Technology Assessment, grant HS 05616. Ann Arbor: Survey Research Center, University of Michigan.
- Converse, Jean M., and Stanley Presser. 1986. *Survey Questions: Handcrafting the Standardized Questionnaire: Quantitative Applications in the Social Sciences*. Beverly Hills, Calif.: Sage.
- DeMaio, Theresa J., ed. 1983. "Approaches to Developing Questionnaires." Statistical Policy Working Paper 10. Washington, D.C.: Office of Management and Budget.
- Fowler, Floyd J. 1989. "Coding Behavior in Pretests to Identify Unclear Questions." In *Conference Proceedings: Health Survey Research Methods*, ed. Floyd J. Fowler, pp. 9–12. Washington, D.C.: National Center for Health Services Research.
- Fowler, Floyd J., and Thomas W. Mangione. 1990. *Standardized Survey Interviewing*. Newbury Park, Calif.: Sage.
- Jabine, Thomas B., Miron L. Straf, and Judith M. Tanur et al., eds. 1984. *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*. Washington, D.C.: National Academic Press.
- Lessler, Judith T. 1987. *Use of Laboratory Methods and Cognitive Science for the Design and Testing of Questionnaires*. Stockholm: Statistics Sweden.
- Morton-Williams, Jean. 1979. "The Use of 'Verbal Interaction Coding' for Evaluating a Questionnaire." *Quality and Quantity* 13:59–75.
- Morton-Williams, Jean, and Wendy Sykes. 1984. "The Use of Interaction Coding and Follow-up Interviews to Investigate Comprehension of Survey Questions." *Journal of the Market Research Society* 26:109–27.
- Oksenberg, Lois, Charles F. Cannell, and Graham Kalton. 1991. "New Strategies for Testing Survey Questions." *Journal of Official Statistics* 7:349–65.
- Payne, Stanley L. 1951. *The Art of Asking Questions*. Princeton, N.J.: Princeton University Press.
- Presser, Stanley. 1989. "Pretesting: A Neglected Aspect of Survey Research." In *Conference Proceedings: Health Survey Research Methods*, ed. Floyd J. Fowler, pp. 35–38. Washington, D.C.: National Center for Health Services Research.
- Rasinski, Kenneth A. 1989. "The Effect of Question Wording on Public Support for Government Spending." *Public Opinion Quarterly* 53:388–94.
- Royston, Patricia N. 1989. "Using Intensive Interviews to Evaluate Questions." In *Conference Proceedings: Health Survey Research Methods*, ed. Floyd J. Fowler, pp. 3–8. Washington, D.C.: National Center for Health Services Research.
- Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys*. New York: Academic Press.
- Willis, Gordon B., Patricia N. Royston, and Deborah Bercini. 1989. "Problems with Survey Questions Revealed by Cognitively Based Interviews." In *Proceedings, Fifth Annual Research Conference*, pp. 345–60. Washington, D.C.: Bureau of the Census.