SELECTING THE QUESTIONS TO BE ASKED

Norman M. Bradburn, National Opinion Research Center and
Graduate School of Business of the University of Chicago

On her deathbed, Gertrude Stein is reported to have uttered these memorable words: "What is the answer? What is the answer? Ah, what is the question?" And that is precisely the state I am in when asked to talk about the problem of selecting questions for a questionnaire. What are the questions about questions that should be asked?

I would like to share with you today some of my reflections on the extremely vexing problem of selecting the questions to be used in a survey research questionnaire. My remarks are directed specifically at the problems of choosing questions relating to social-psychological variables, which are being increasingly used in survey work concerning social programs. I believe, however, that many of the considerations relevant to the selection of questions on social-psychological variables will be pertinent to the choice of questions on other types of variables.

My reflections have led me to focus upon the difficulties involved in formulating adequate decision rules for the selection of social-psychological questions. As I see it, there are five principal obstacles to the development of clear decision rules:

1. A lack of agreement among behavioral scientists about the appropriate social-psychological dependent variables that are relevant to particular social programs;

2. An inadequate conceptualization of those social-psychological variables that are suggested for study;

3. A relative lack of interest in systematic methodological research and survey measurement;

4. The relative underdevelopment of measurement theory in survey work as compared with the sophistication of sampling theory; and

5. The special historical and cultural problems that affect the phraseology of questions.

Let me discuss each of these obstacles briefly. The first two are intimately related to one another and concern the problem of knowing what concepts one wishes to measure. First of all, social psychologists themselves cannot agree on what the relevant variables should be, and second, because of the poor conceptual development of the variables that are suggested, there is not even a reasonable degree of consensus on their relative importance. Let me give as an example some problems that NORC faced in doing surveys connected with the evaluation of Manpower Training Programs. Increasingly, those involved in the evaluation of these programs believe that some kinds of social-psychological factors may be important in determining the effectiveness of training programs, and even in some instances, that particular types of social-psychological changes should be viewed as important outputs of the program in addition to increased skill level. The extension of interest to factors other than narrowly defined economic ones is to be applauded. It is lamentable, however, that behavioral scientists, who appear to have succeeded in convincing economists that there is something more to the world than economic variables, cannot now come forward with better suggestions regarding the social-psychological variables to be studied.

What are the kinds of variables that behavioral scientists might propose as relevant to job training programs? Some would suggest variables concerned with motivation, while others would emphasize the relationship between personal interests (or needs) and the characteristics of the job itself or the work environment; some might stress variables relating to job satisfaction, while others would stress variables relating to managerial or supervisory styles. Still others might view the important variables to be those related to an individual's general disposition toward society or alienation from work. In a general way these differing approaches can be grouped into those approaches concerned with individual differences and with variables that are conceptualized as being within the individual; those approaches that stress variables relating to work organization and the social setting within which jobs are performed; and those approaches that cover both fronts by pointing to the importance of the interaction between the two general types of variables.

It is easy to point out that there is a lack of consensus among behavioral scientists about the importance of particular variables. It is not so easy to explain why such a lack of consensus should exist, or to suggest things that could be done to improve the situation. Although I have no good evidence, I strongly suspect that the situation is fostered by two trends in the behavioral sciences that have been going on for some time. The first is that most research in the behavioral sciences limits itself to establishing the existence of relationships among variables, and gives practically no attention to the assessment of the magnitude of those relationships. If one reads the professional journals in these fields, one has the feeling that psychologists and sociologists worship a god of "statistical significance" and have an unquestioning faith that all things which are statistically significant are equal in importance in the world and that all things which are not statistically significant are totally unimportant. Closely allied with such a mystical belief and aiding it is the fact that the great preponderance of research is carried out under laboratory or other artificial conditions which make its applicability to real world phenomena extremely

limited at best. Neither of these aspects of research is calculated to aid one in deciding what the practical importance of particular variables might be in real life situations. As the demand from those involved in applied behavioral research increases, I suspect that we shall see some changes in these trends. When challenged "to put up or shut up," I find it hard to believe that behavioral scientists will be able to shut up.

Achieving a greater agreement about the important variables to study would help in clarifying decision rules for selecting questions, but it certainly would not go the whole way. Even among those who agree in general about the importance of particular variables, there can be serious division regarding the conceptual status of specific variables. Take, for example, a variable such as job satisfaction. Is this a single-dimensional, two-dimensional, or n-dimensional variable? Is job dissatisfaction the opposite end of a single scale, with job satisfaction at the other end, or should job satisfaction be conceptualized as consisting of a single general factor with some number (usually unspecified) of specific factors? Since job satisfaction is probably the area in which the greatest amount of research has been done, much of which has been conducted with fairly direct applied interest, one cannot be very optimistic about the chances for early clarification.

If one looks at notions of work motivation, the picture is even murkier. Is a motive conceived as a generalized energizer or as a "push" toward particular goals or activities? Are motives related to kinds of activities or to the value of certain outcomes? How do motives differ from occupational values? The ways in which these problems of conceptualization are resolved will have significant implications on how the variables are viewed in relation to the particular program being evaluated, and therefore on the kinds of questions selected. As with the problem of agreement on the relevant variables, one can only hope that the greater need for conceptual clarity will force behavioral scientists to think through more thoroughly the nature of their concepts and to do the necessary work to resolve the difficulties.

Let me now turn to two other related difficulties in the current status of the development of survey research methodology. These involve the apparent lack of interest in systematic methodological research and the underdevelopment of measurement theory in survey work as compared with the development of sampling theory. It seems a safe generalization that in the field of survey research, methodological research has a relatively low priority. This is not to say that there is no methodological research being conducted, but rather that the research that is done tends to be fragmented, local, unpublished, and usually specific to particular studies. The bulk of methodological research appears to be done for the internal benefit of large organizations, such as the Census Bureau, the Survey Research Center at the University of Michigan, and NORC. I expect that other academically based survey organizations, many commercial and market research firms, and individual scholars scattered around the country have also done important methodological work centered on their particular concerns. However, for the most part this work does not get published. With very few exceptions, systematic methodological work is notable by its absence.

It is easy to see how this lack of interest is perpetuated. There is comparatively little professional payoff (at least in the behavioral sciences) for good methodological work in the area of variable measurement. Such payoff currently goes to those who work on techniques of data analysis and in sampling theory. It is more difficult to discover why the reward system should operate in this fashion, because one would think that the value of sophisticated analytic techniques would be negated by the poor quality of the data being analyzed. Behavioral scientists appear to have a curiously ambivalent attitude regarding their data. Sometimes they take the stance that they know perfectly well that most of the data they use are of very poor quality and have a high amount of error, but they believe that sophisticated analytic techniques will enable truth to shine through all the data noise. On the other hand, at times they write and publish as if they were blissfully unaware that any serious measurement problems existed in the data and as if they need to be concerned only with sampling error, and sometimes not even with that.

How does one explain this comparative lack of interest in response errors as compared with the fairly sophisticated development of psychometric models in the measurement of educational achievement and individual differences in abilities? I suspect, although I have no evidence, that the difference lies in the uses to which data in social research have been put, compared with the ways in which data are used in educational systems and in personnel selection and placement. For the most part, survey data in the behavioral sciences have not been used to make important decisions concerning people's lives. There is every indication, however, that survey data will play an increasingly important role for social-planning purposes and in the evaluation of social programs. In addition, the allocation of large sums of money is now and will increasingly in the future be influenced by the results of sample surveys. Unfortunately, I believe that we are presently at a state where the quality of most survey data is too poor to support the uses to which it is put. We must work to improve very significantly the quality of data that we are collecting.

There are a few signs that things are beginning to change for the better. One of the most promising is the publication by the Institute for Social Research at the University of Michigan of two volumes that attempt to survey the state of measurement in the areas of political attitudes[1] and occupational attitudes and characteristics.[1] John Robinson and his colleagues bring together

in one place, to my knowledge for the first time, many different measures of the same variable, and attempt to review systematically the adequacy and usefulness of each measure. In many respects these volumes parallel those that have existed for years in the area of ability and achievement tests. It is perhaps an eloquent comment on the state of concern for methodology in survey work that such volumes should be only just now forthcoming. I trust that this is the beginning of a series which will be continuously improved and updated.

The authors of these volumes have used three groups of criteria to evaluate the scales reported. These are: (1) criteria relating to item construction, such as the sampling of relevant content, item analyses, and adequacy of question wording; (2) response set criteria, that is, assessing the attention paid to problems of response bias, acquiescence set, etc.; and (3) psychometric criteria, such as measures of reliability, normative information on the scale, and the ability of the scale to discriminate between groups known to differ in the dimension of concern. The fact that these sets of criteria have been applied consistently to the evaluation of the scales presented in these volumes adds immeasurably to their usefulness. The overwhelming impression that one receives from reading these volumes is that even in areas which have been fairly well studied, such as the measures of job satisfaction, the level of scale development work is surprisingly low and the information necessary for adequate evaluation of many measures is simply not available. I hope that the publication of these studies in systematic form will help rectify this situation.

Finally, I would like to mention a set of problems that are extremely disturbing to anyone seriously interested in the methodology of question asking--particularly disturbing because they seem so intractable. These difficulties stem from changes in linguistic usage across time and variations in usage among different subgroups within the same population or across different populations. In short, these are the vexing problems of the comparability of question meaning to the respondents at different points in time or at different points in space. Shifts in the use of language over time is a particular problem if one is interested in monitoring social change, such as changes in attitudes toward certain types of programs (for example, social security or social welfare programs), or is concerned with measuring changes in racial attitudes. To give just one example of the kinds of shifts that occur in word usage, in a study of racial attitudes in 1950 NORC asked white residents of neighborhoods in Chicago the following question: "Do you approve or disapprove of white and colored children being in the same schools together?" In 1964 we were asking the question in this form: "Do you think white students and Negro students should go to the same schools or to separate schools?" I expect that in the near future we shall be asking the question in terms of "Do you think white students and black students should go to the same or separate schools?" These alterations in wording are small and may be inconsequential adjustments to shifting usage. But on the other hand, we don't really know what are the effects of the changes. We all know that slight variations in the wording of a question may bring about relatively large shifts in the distribution of responses, but we have no systematic data that would allow us to approximate how much of the change in distributions is due to alterations in question wording and how much is due to real changes in opinion.

It is perhaps fortunate that we do not have much that can pass for trend data in social or political attitudes, since the problem of change in question wording over time has not really been seriously faced. However, if we should begin to rely on surveys for significant social indicator data, much more serious consideration will have to be given to this problem.

The problem of wording differences is not confined to changes over time but is also omnipresent in cross-sectional research. We have taken it more or less as a canon of faith in survey research that all respondents--all those who speak English at least--should be asked the questions in exactly the same wording, regardless of their educational level. The result of this article of faith is that questions addressed to nationwide samples are couched in a vocabulary that is presumed to be understandable by even poorly educated respondents. Such a presumption, however, may have little basis in fact for we know little about the way in which poorly educated or minority-group respondents interpret questions worded in standard, albeit simplified, English. Many of us in survey research feel that there is probably considerable loss of information when the identical question wording is used for both middle-class and extremely poor or minority-group respondents. But we have not found any usable way to alter question wording so that it becomes appropriate to the characteristics of the respondent. On the other hand, I don't think we have tried very hard either.

Recently the Social Science Research Council's Committee on Sociolinguistics sponsored a conference on "Language as Obstacle and as Data in Sociological Research." Participants in this conference pointed out that we need to be concerned not only with the problem of the meaning of words and sentences in interview schedules, but we should also pay some attention to the context within which interrogative sentences are embedded. The SSRC group called for more attention to what they termed "the ethnography of asking questions." In his report on this conference Allen Grimshaw noted:

> We simply do not know how to phrase questions that will be meaningful to random samples of diversified populations. We suspect that fixed-choice questions should never be used in comparative studies. Since those who use them have not systematically examined the possibly resulting biases, however, we have no way of estimating the magnitude or direction of errors

that are thereby introduced. . . .

> Social scientists must ask themselves some serious questions about how wisely they ask questions of their research subjects. Otherwise, continuing refinements in quantitative analysis of data will produce only spurious or at best marginal increments of socially and sociologically relevant data.[2]

This is a view that I believe deserves very serious consideration by all those who are active in the area of survey research.

From the foregoing discussion of five obstacles to the delineation of precise decision rules for selecting questions, we can now point to some rules of thumb for use in asking questions about what questions to select. I would summarize these rules of thumb as follows:

1. What is the theoretical relevance to the problem at hand of the social-psychological variables that I select for study? If I cannot specify what these variables are and at least what their theoretical relation is to the phenomena I am studying, then I should abandon the effort to measure them.

2. From the best information I can get on operational measures of these variables, what measure best meets the criteria of good item construction and least susceptibility to response biases and has the best psychometric properties? It may turn out at this point that there are no scales which meet the minimum standards I have set for my research, and I must either abandon the effort to measure the variables or embark on a side excursion in developing new measures.

3. Finally, what are the characteristics of the population that I am surveying, particularly with regard to degrees of heterogeneity which might require different forms of questions for different segments of the population to be studied? While I believe it unlikely that one will find any measure that will have alternate forms of questions for differing subgroups of the population, I believe that the researcher should seriously ask himself whether he should not devote some of his precious research time to investigating the potential biasing effects of using the same question form for all respondents.

I have no illusions that the practice of survey research comes anywhere near to approximating the ideal toward which we strive. I do, however, feel strongly that we must make very substantial improvements in our measurement standards if we are to fulfill the promise that survey research methodology has made to those who are engaged in social research.

FOOTNOTES

[1]John P. Robinson, Jerrold G. Rusk, and Kendra B. Head, Measures of Political Attitudes (Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan, September, 1968); John P. Robinson, Robert Athanasiou, and Kendra B. Head, Measures of Occupational Attitudes and Occupational Characteristics (Appendix A to Measures of Political Attitudes) (Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan, February, 1969).

[2]Allen D. Grimshaw, "Language as Obstacle and as Data in Sociological Research," Items, v. 23, no. 2 (June, 1969), p. 21.