

1. Introduction

This paper describes the results of a project which examined respondent performance in completing a version of the U.S. decennial Census long form. The results are based upon a series of observational studies conducted under a contract with the U.S. Census Bureau (Contact No. 50-YABC-6-66032).

Two types of observational studies were conducted, one-on-one observations and group observations. The one-on-one studies provided and opportunity for observing the entire process of completing the form including the respondent's overall approach, understanding and misunderstanding of the conventions used in the form, the problems that respondents had with moving through the form, and their problems with individual items. This paper describes the results of the one-on-one observations.

The Census long form is mailed to approximately 25 percent of U.S. households. Each household receives a form and a separate instruction booklet. As well as the basic enumeration questions, the long form contains a series of detailed person and housing questions. Not all forms are returned and many of those that are returned have errors. Previous studies (Rothwell, 1983 and 1985) had reveal a variety of problems that respondents had when completing the form. These included difficulties understanding and answering particular questions, carryover effects, and position effects. In addition, Behavioral Analysis Studies by the Bureau reveal that respondents who began filling out the form generally did return it whereas those who failed to begin did not return it. These studies underscored the need for a "user friendly" form.

The one-on-one observational technique that we used was the collection of response protocols (Nisbett and Wilson, 1977) in which respondents are requested to give verbal reports of their mental processes as they complete a task. The reports are recorded and subsequently analyzed.

Respondents were given a pencil, a blank form, and the Instruction Guide. Some observations were conducted in the respondent's home; others in an office, classroom, or library. In every case, the observer chose a place where there was plenty of space for spreading out the form, plenty of light, and few distractions. Thus, the respondents were freed from the task of finding a suitable setting for completion of the form, which in some cases made the task easier than it would have been if the form arrived in the mail at their homes. However, in other cases, this set-up added difficulty to the response task. Respondents who normally would have had and used household records could not do so. Those who would have called upon other household members for assistance or who would have taken breaks from a tiring and difficult task were prevented from doing so. The observers attempted to give only the help that was necessary to assure that the respondent could continue through the form and answer a reasonable number of questions in the allotted time.

The respondent was asked to report his thoughts as he completed the questionnaire. In addition, if the respondent gave few reports, he was asked why he chose certain answers or did certain things.

The one-on-one respondents ranged in age from 20 to 68 years, included both men and women; Blacks, Whites, and Indians; and people with doctorates as well as people who had not completed high school. They varied greatly in what we have come to call "forms literacy." We consider forms literacy to describe the degree to which the respondent is cognizant of the goals, structure, conventions, and semantics of forms. Some of our respondents were professionals engaged in the business of survey data collection and, therefore, were very familiar with the design and conventions of forms. Others were so unfamiliar with the conventions of forms that they could not determine how to begin the task.

During these interviews, it became quickly obvious that the response task had two components, 1) the inherent recall and reporting tasks and 2) the task of attempting to understand the conventions of the form. All respondents were required to understand questions, recall information, and mark the correct response on the form. Those who were not forms literate were faced with an additional task--understanding the conventions of a self-administered form. For some respondents this was bewildering task with multiple novel elements--the filling of circles to indicate response choices, the interpretation of complex conditional instructions, the need to understand and interpret unfamiliar terms and concepts, and the requirement for dealing with inconsistent conventions within the form combined with the need to recall information, make calculations, and select responses. The same respondents who were not well-versed in the completion of forms were, sometimes, the same respondents who had difficulty seeing due to poor or absent corrective lenses. This same group lacked confidence in their ability to make sense of the task and when they, for example, missed a skip pattern that caused them to answer inappropriate and, as a consequence, nonsensical questions, they became even more bewildered and naively continued answering rather than searching for an error. This appeared, in some cases, to cause them to decide that the Census Bureau was only interested in superficial responses.

All respondents, regardless of level of form literacy, had problems with the form. Some of those who had a high degree of forms literacy made mistakes because of an apparent assumption of proficiency. These people would often skim the questions rather than reading them in detail thus missing skip patterns or pertinent instructions. When completing the second or later sets of detailed person questions some tended to not read the questions a second time thereby missing qualifiers and answering questions incorrectly for subsequent persons.

2. Response Process

A respondent goes through several steps when answering survey questions. Oksenberg and Cannell (1977) presented a model of the question-answering process for an interview survey.

According to their model the question-answering process consists of:

1. Comprehension of the question in which the respondent tries to understand what is being asked.
2. A cognitive processing step in which the respondent makes some decisions as to the information needed, attempts to recall information, and then formulates a response based upon his recall.
3. An evaluation step in which the respondent judges the accuracy of the response.
4. A second evaluation step in which the respondent may evaluate his response in terms of other goals. For example, the respondent may be hesitant to report certain attitudes or behaviors.
5. As a result of this evaluation step, the respondent may decide to give a response that he or she judges to be accurate or may modify the response based upon other considerations.
6. Some of the other considerations that affect response are cues from the interviewer, the respondent's values and beliefs, cues from other questions, and so on.
7. The respondent may then give a response that is inaccurate in some way. The respondent may or may not be aware of these other cues produced a modification in the response.

A respondent in a self-enumeration survey must perform a number of additional activities in order to provide the responses. Based upon the results of our observations, we have developed a model for the process of responding to a self-administered questionnaire. This is shown in Figure 1.

1. The respondent must read and understand the instructions.
2. The respondent must examine and understand the structure of the form. This requires completing several subtasks, including finding the questions, comprehending the order of questions, understanding the relationship between blocks of information and sections of the questionnaire. The Census Long form had a fairly complex structure that remained a mystery to several of the respondents in our study.
3. The respondent must find the first question. This may require him to read and interpret conditional statements if some people are to complete one section and not others.
4. Once the respondent locates the first (and subsequent questions) he must decide if the question applies to him. This will require him to interpret

conditionals, skip instructions, and the question response structure.

If the respondent concludes that the question does apply, he then undertakes steps 5-9. Otherwise he must go to step 11.

- 5-9. Steps 5-9 correspond to the steps in the Oksenberg and Cannell model.
10. If the respondent decides that the question does not apply to him, he must then determine what to mark on the questionnaire. Conventions vary. Some forms provide space for marking "not applicable"; others require that the respondent merely skip the question without making any marks.
11. Indicating the Answer. The respondent may be required to write the answer on lines or in special blocks. In other cases, the respondent must choose one of a group of responses. This again requires examining the question/response structure, reading and understanding the categories, and finally marking the chosen response.

In other situations, the respondent is required to code the response. This requires interpreting codes, choosing the correct one and entering the codes.

12. Find the Next Question. Sometimes this requires the respondent to read and interpret skip instructions or to interpret conditional statements.
13. Determine When All Questions Are Answered. This task is not always easy. In some cases the form is constructed with multiple skip patterns, and large blocks of questions do not apply. In the case of the Census long form, person pages are provided for 7 different household members. Thus, in small households, much of the form is to be left blank.
14. Finally, in order to complete the response task, the respondent must return the form. This will require packaging and mailing. Usually postage will be provided.

3. Problems Experienced by Respondents

There are a number of ways in which the problems that respondents had could be classified. We have developed a hierarchy of problems that includes three levels -- problems the respondents had getting started, problems moving through the form, and problems answering specific questions.

Difficulty getting started is the most serious type of problem. It represents failure to successfully complete Tasks 1 - 3 described in Figure 1. A respondent who can not understand the overall nature of the task may never begin and, if he does begin, may make such serious errors that his information is unusable.

Once the respondent begins the task, he must solve the problems inherent in moving through the form so that he answers the questions that are appropriate to his situation. This requires successful completion of Task 2, Task 4, Task

11, and Task 12. If this is not mastered, large blocks of data may be missing or inappropriately included.

Finally, the respondent must adequately answer the individual questions (Tasks 5 - 10). However, failure to understand a particular question is not as serious as failure to master the superordinate tasks because errors at the question level will not necessarily affect subsequent questions.

In our study, the very skilled respondents took some time to look at both the form and the instructions and they quickly grasped what each contained. These people generally would briefly examine the instructions and then decide to proceed with the completion of the form. Knowing that they possibly "should" read the instructions, several asked whether they were really supposed to do what they would do if the observer were not there. When assured that they were, they put the instructions aside and began with the form, never or seldom to return. These respondents focused most of their attention on answering the questions (Tasks 5-10) and did not have to devote much effort to the other tasks.

Other respondents had great difficulty comprehending this overall organization; some never understood it.

Problems included:

1. Difficulty understanding the distinction between the form and the instructions.
2. Difficulty determining where the form began.
3. Failure to understand that people are to be listed by name.
4. Failure to understand the one person one column format.

One respondent who had difficulty getting started but who was finally able to make progress was a 46-year-old high school graduate enrolled in adult education classes at a local community college. He spent quite some time examining the instructions, the form, and trying to decide how he was to begin. He expected to find an instruction for each question and was disturbed when he could not. He read very carefully and, for a time, got caught in a loop in which he went back and forth between the form and the instructions trying to decide what to do.

Another serious problem in getting started was illustrated by a 63-year-old woman who was not a high school graduate. This woman lived with her grandson. The form has a "person column" for each member of the household. She listed both herself and her grandson in the first column. If the answers were the same for each of them, she marked one answer. If they were different she marked 2 answers.

A 27-year-old woman with some college education listed the eight people in her household by their relationship to herself rather than by name, i.e., mother-in-law, father-in-law, ..., husband, myself.

Another woman who was a 40-year-old high school graduate examined both the form and the Instruction Guide and decided that she had two things--a question sheet and an answer sheet. She tried to read as the questions the material in the instructions labeled "Instructions for

Questions 2 through 7." She could not determine what to do until instructed by the observer.

These difficulties and failures in beginning the response task are of major significance to the conduct of the Census. The foremost goal of the Census is population coverage. It is likely that several of these respondents would not have completed the form at all with out the help of the observer. New design efforts must, therefore, place major emphasis on assisting those respondents for whom the understanding the overall nature of the task will be difficult.

Understanding the Structure of and Movement Through the Form

Those who were not forms literate had difficulties moving through the form. The form presumes that the respondent will understand that a single person is to be listed in each of the person columns and that the questions should only be answered for one person, and, unless instructed, there will be only one answer for a question. The form is inconsistent in what it requires the respondent to do if a question does not apply. In some cases the respondent is to skip the question; in others, the respondent is to indicate that the question does not apply.

The structure of the form assumes that the respondent understands conditionals. Several questions are introduced by the phrase "If you..", sometimes written in bold as part of the question, sometimes written in italics above the question. Another assumption is that the respondent will recognize partitioned questions and answer all parts keeping in mind the introductory phrases. In addition, in some cases, new questions are introduced in the answer categories. For example, Question 15a asks where the person lived 5 years ago, and the first answer concerns when the person was born; a question on plumbing facilities has qualifying statements included in the answer categories which, in essence, introduce a new question on whether or not the plumbing facilities are shared by more than one household.

Problems that the respondents had in understanding the structure of the form and moving through included three types of problems:

1. Failure to grasp the basic structure of the form including the one person one column convention, the need to maintain a correspondence between the person columns and person pages, and the fact that the person questions should be answered for one person at a time.
2. Failure to understand conditionals.
3. Difficulty seeing, understanding and/or following skip patterns.

Problems Answering the Questions.

A respondent who has begun the task and is moving from place to place in the form may still have difficulty in with the individual questions. We identified three types of problems:

1. Overall problems with the question response structure.
2. Problems coding the responses.

3. Problems understanding specific words or phrases in the questions

The first two of these problems were more prevalent among those were low on forms literacy because they had little or no prior experience to guide them in interpreting what was required. These people were also likely to have more difficulty with specific words and phrases because they were less literate in general.

Many detailed examples of these problems are given in the report from the study (Holt and Lessler, 1987). The results from this study have a more general implication to understanding the nature of the survey measurement process. First, these results suggested a general process model for the tasks that respondents must complete when responding to a self-enumeration survey. Second, understanding the response process will allow us to identify the types of survey errors that will result from difficulties that respondents have with the components of the response process. Unit nonresponse, item nonresponse, and measurement errors are directly related to certain components of the task. Failure to begin the task results in unit nonresponse in a sample survey and undercoverage in the Census. Problems understanding the structure and movement through the form results in item nonresponse and measurement errors. Problems answering specific questions results in measurement errors and item nonresponse. Gaining a detailed understanding of the response process will enhance our efforts to improve the survey measurement process.

Responding to a self-enumeration survey is a complex task for many people. It requires a number of skills that they may not have. There will be major differences between subgroups of people in the accuracy and completeness of information because of their differing abilities to complete the required tasks. This can introduce serious errors in our conclusions particularly when different subgroups are compared. Gathering response protocols provided us insight into these problems and suggested potential solutions. We suggest that this is a good method for gathering information that will allow us to direct our subsequent efforts to improve the survey measurement process.

References

- Holt, Mimi and Judith T. Lessler (1987). An Evaluation of the Long Form Questionnaire for the 1986 Census of Central Los Angeles County. FINAL REPORT: QUESTIONNAIRE DESIGN PROJECT. Contract No. 50-YABC-6-66032.
- Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Rothwell, Naomi, 1983. "New Ways of Learning How How to Improve Self-Enumeration Questionnaires: A Demonstration Project." Bureau of the Census Report, Sept. 1983.
- Rothwell, Naomi, 1985. "Laboratory and Field Response Research Studies of the 1980 Census of the Population in the United States." *Journal of Official Statistics*, Vol. 1, No. 2: 137-158.

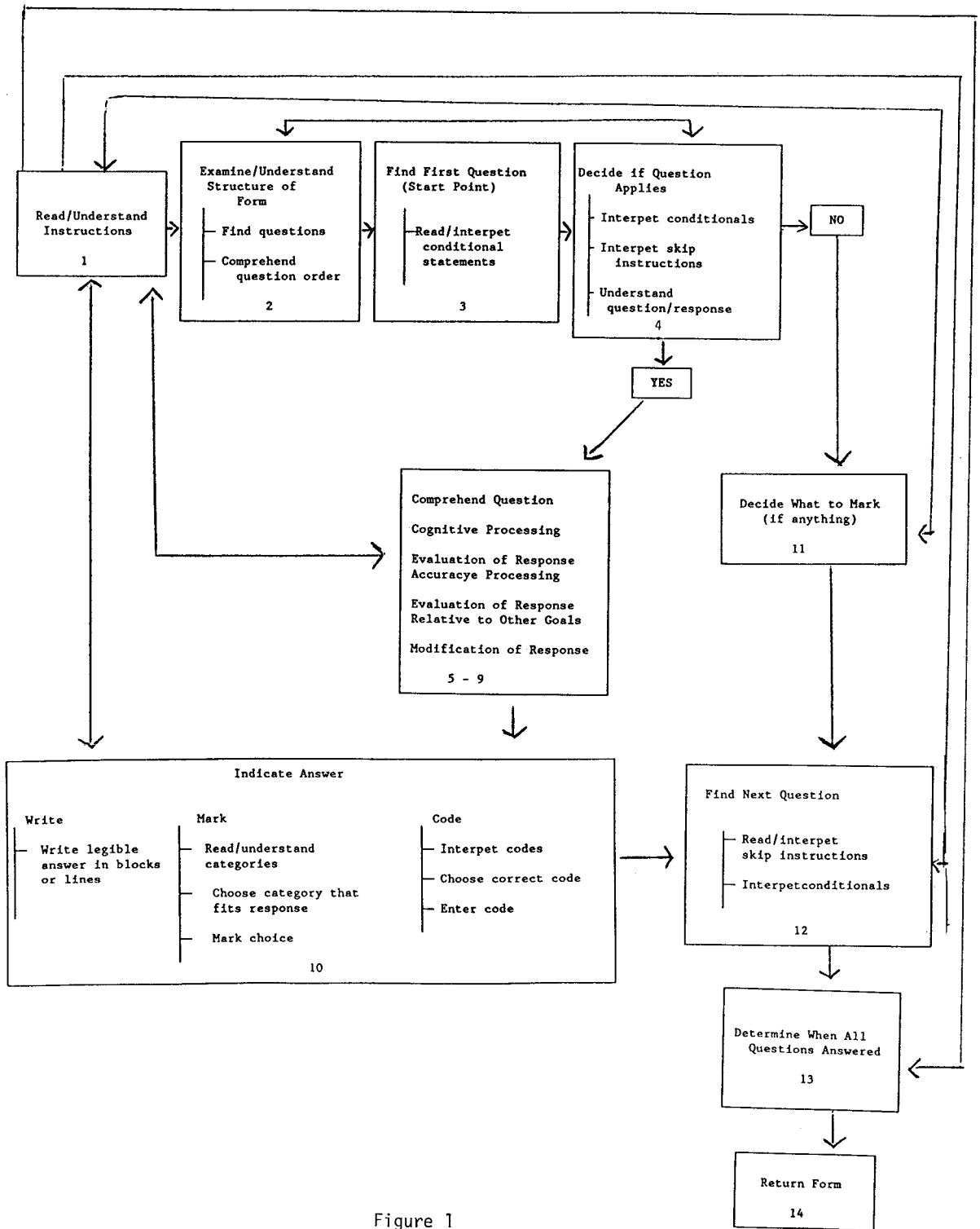


Figure 1