

Theresa J. DeMaio, Elizabeth A. Martin, and Elizabeth Page Sigman
U.S. Bureau of the Census¹

Introduction

The decennial census is conducted largely by self-enumeration. Therefore, the census questionnaire critically affects whether or not people participate in the census, and the completeness and accuracy of the information they provide. Questions, instructions, and response categories must be clear both conceptually and graphically.

The Census Bureau is engaged in a program of questionnaire design research that aims to improve response rates, and data quality by looking at the task of answering the census questionnaire from the respondent's point of view.

This paper presents initial results from this research. We took as a starting point the long form from the 1986 Census of Central Los Angeles County. To understand the problems that respondents had in completing the census long form, Census Bureau experts observed respondents as they completed forms. Each session consisted of a respondent and an observer who carefully watched whether or not the respondent read instructions, in what order the items were completed, etc., and what problems were encountered. After the form was completed, the observer debriefed the respondent to elicit information about problems and misinterpretations. Observers asked respondents what they had in mind while answering items, how they understood the concepts in the question, whether they recognized skip instructions, and other probing questions. These sessions provided a wealth of information not available by simply reviewing completed forms and examining the errors that appear on them (see DeMaio, 1986b; DeMaio, 1986c; and Holt and Lessler, 1987 for detailed summaries of problems with the 1986 census long form).

Based on these results, inspection of the form, and consultation with experts both within and outside the Census Bureau, we revised the census long form. The revisions we made are described in the sections below.

To evaluate the effectiveness of our revisions, we conducted some small-scale quantitative tests of our revised questionnaire. These were split-panel classroom tests in which half of the participants were randomly assigned the revised long form, and the other half the 1986 test census long form. Participants in each session first completed a census form, and then were given a debriefing questionnaire which tapped their understanding of concepts and terminology used in the questions.

A total of 515 people participated in tests which were conducted in April 1987. Participants included people aged 18 to 80, members of different racial and ethnic groups, and people with various levels of education. Minority racial and ethnic populations with relatively little education were overrepresented. The participants were recruited by Census Bureau regional office staff who work with local community organizations. Using

their contacts, the staff organized sessions as part of regular activities, or organized special sessions with employees, group members, community service participants, etc. The participants in these sessions do not constitute a representative sample, but provide us with enough cases to make some broad statements about statistical differences in reporting on the two forms.

In the next three sections of this paper, we discuss problems related to the overall structure of the form, those related to response formats, and conceptual and question wording problems. We describe our revisions to the questionnaire and present results of our tests.

Problems Related to the Structure of the Form

In our initial one-on-one observations, we discovered that many respondents had difficulty understanding and following the form. A few people answered the example questions in the instruction guide before they even noticed the census form. Other people repeated information about the same person more than once because they didn't realize that the same set of questions was to be answered for each household member. Three characteristics of the form contribute to these problems. First, the sequence of sections, and the order of questions within sections, do not make sense from the point of view of respondents. Second, routing instructions do not explain the structure of the form. Third, terminology for referring to household members is not consistent.

The structure of the long form is imposed by the need for quick processing of 100 percent information from all households and persons. The operational requirements of census-taking are accommodated by putting all 100 percent person and housing questions on a single large foldout page, which is followed by the sample housing and person questions. Thus, the long form involves a fairly complex sequence of tasks: respondents must first list all household members, fill a matrix of information about each person, answer coverage and housing questions, and finally answer additional questions about each person listed.

This awkward order is confusing because the sequence of tasks is not clearly labelled. Nor are questions numbered sequentially--the housing items are preceded by an H (H1 to H29), and are placed between two sections of person items (items 1 to 7 precede the housing items, and items 8 to 31 follow them). The lack of clear structure can result in major respondent errors. Some respondents do not know where to start, and skip Question 1 (the listing of household members) entirely. Some respondents do not see that the flap is to be opened, so they skip page 2, which contains the 100 percent questions for each person in the household. These questions are presented in a matrix format, with the questions listed down the side of the page and the names of household members across the top. Some people do not understand

this matrix format and make mistakes. Finally, there is no clear indication when or where to stop filling out the form. As a result, some respondents report duplicate sample person information. Many people waste time flipping through the form trying to figure out where to go next, and do not ever find the space on the back page to provide their name and telephone number.

Another problem is that the terminology used to identify household members is not consistent throughout and is not reinforced. The original listing in Question 1 does not label or number the lines on which names are to be entered by the respondent. At the second stage, while completing the matrix of 100 percent information, the boxes in which names are to be written are labelled "Person 1," "Person 2," etc. One of the subsequent housing items refers to these labels. However, upon confronting these labels after answering numerous questions on unrelated topics, it is not certain that respondents associate the labels with the proper persons from the matrix listing. This may cause mismatches between the various sections of person data, and may encourage improper definitions of household membership.

We attempted to eliminate problems by revising the form. First, we tried to make the form easier to follow by clearly labelling the sequence of tasks ("Step 1. Make a list of the people who live here," and so on). Second, we numbered the questions consecutively from the beginning through the last person question. The "H" prefix for housing questions was dropped. Third, we clarified references to persons. Instead of referring to "this person" or to "Person 1," "Person 2," etc., we consistently referred to "Person A," "Person B," and so on. Letters were used instead of numbers, since numbers are used to label questions and in many other ways. All population items refer to an explicit person either in the question (for sample items) or in the responses (for 100 percent items). These letters are introduced in Question 1 and maintained throughout the form. Finally, we added instructions to make it clearer to respondents when they are finished.

Table 1 presents a general summary of the data necessary to evaluate the effects of our changes. A total of 515 forms were filled during the classroom experiments--256 1986 forms and 259 revised forms. The households enumerated in the two groups were of approximately equal size--the 1986 forms contained an average of 3.5 persons, while the revised forms contained an average of 3.7 persons.² Since the average household size is similar, one might expect data would be given for similar numbers of persons on both forms. However, as Table 1 shows, data were given for substantially higher number of people with the revised than with the 1986 form (743 vs. 703). This suggests that respondents who filled out the revised form provided more complete data for all household members.

Other data also suggest that revised forms were filled more completely. We had observed that some people did not know where to begin filling out the form and skipped the household

listing. In our revision, we sought to clarify instructions and help respondents to begin the form. This effort appears to have been successful. Only 7 percent of respondents to the revised form left the household listing blank, while 28 percent left it blank on the 1986 form.³ We believe that the simplified instructions on the front cover and flap are responsible for this improvement in response.

Table 1 also shows that respondents to the revised form were significantly more likely to begin the sample person section, and to finish the sample person section once they started it. Five percent of the 1986 forms contained no sample data, compared to less than 2 percent of the revised forms. In addition, 16 percent of the 1986 forms lacked sample data for one or more household members, compared to 9 percent of the revised forms. Participants in the classroom experiments were allowed only 55 minutes to fill out the census form. Respondents to the 1986 form may have provided more complete data if they had been allowed unlimited time. However, these data suggest that the revised form was easier for respondents to complete, given a limited amount of time.

We also thought our revisions would reduce the likelihood of leaving the person matrix completely blank. We hoped to reduce reporting of duplicate person information, and to decrease the incidence of reordering persons between the 100 percent and sample person sections. Results indicate these problems occurred infrequently on both forms. This suggests that the two-part presentation of the person items may not be as difficult as we originally thought.

Response formats

We found a number of problems with the organization of response categories and formats for entering responses. Some questions present respondents with long, complex lists of unordered response categories. Sometimes the visual organization of the categories makes it appear that two circles are to be filled in, when respondents are only supposed to fill in one. In other cases, familiar information (e.g., an address or a dollar amount) is to be entered in an unconventional format. We simplified and reorganized response categories and formats to try to make the response task easier and clearer.

Perhaps the most complex and error-prone response format is machine-readable marking for age and year of birth. The need to quickly capture information about age for the entire population requires that respondents provide age information in a machine-readable format. The task of FOSDIC age-marking is one of the more complicated things that respondents are asked to do in filling out their forms. For this reason, among others, the census attempts to obtain the information in more than one way. The form obtains written and machine-readable reports of age as well as written and machine-readable reports of year of birth. The "multiple measures" solution to age reporting problems is not totally successful, however, since it introduces the possibility of inconsistencies in reports of age and year of birth.

Previous research has provided some evidence of consistency problems. A review of census forms from the 1985 Test Census of Tampa, Florida showed that only 59 percent of the cases contained internally consistent responses for all items (Spencer, 1985). The remaining cases had varying degrees of missing data or inconsistency, and 6 percent of the errors observed could not be fixed without a clerical review.

These errors and inconsistencies are for the most part ignored in data processing; in 1980, only the coded year of birth was used to determine a person's age. Written entries were examined only if the coded year of birth was completely blank. The written entries were never used to verify the accuracy of coded year of birth, primarily because of cost and time considerations (Spencer, 1986). For this reason, it is extremely important that the data be entered correctly by respondents.

We believe that some inconsistencies may be caused by ambiguities in the wording of the age question, which asks for "age at last birthday." This wording suggests that something other than current age is intended, and can be interpreted to mean, "age before last birthday," or some respondents may read it as "age at next birthday." These speculations are consistent with research in the 1985 test census (Spencer, 1985) which found that in about 3.6 percent of person-records from a sample of Tampa test census forms, age and year of birth were inconsistent by exactly 1 year.

The layout of the age item is confusing. There are three rather large spaces for writing age (see Figure 1a). In our observations, respondents did many creative things in answering this item. Quite commonly, people don't realize that the first column should be left empty, and put two-digit ages in the first and second, rather than the second and third columns. Other people see the wide spaces and write two-digit ages in the first column. Some

a. Age			b. Year of birth		
0	0	0	8	0	0
1	1	1	9	1	1
2	2		2	2	
3	3		3	3	

Figure 1a.

people put the month and day of their birth in the second and third columns. Many people leave the FOSDIC circles blank because they just don't know what to do with them. Some people figure out what to do by looking at the instruction guide (some people fill out the example in the instruction guide as well!). The correspondence between the numbers above and the circles below is not always clear. For example, a person with a write-in age of 19 was FOSDIC-coded as 109. Apparently the respondent figured out halfway through the task what she was

supposed to do, but did not go back and make her entries consistent.

The year of birth item also has design problems associated with the machine-readable reporting task. The current format of this item is not a familiar or obvious format at all (see Figure 1a). There are spaces for four digits for year of birth, the first of which is already filled in with a "1." The four columns beneath these spaces have varying FOSDIC marking patterns. The most questionable aspect of the marking patterns is the blackened circle with no number beside it under the written-in "1." It is not clear what this confusing feature is meant to accomplish. Another somewhat inconsistent feature of the current format is that the second box has underneath it only two FOSDIC circles ("8" and "9") which are not aligned with the corresponding circles for the last two digits.

We revised the age and year of birth items problems in several ways (see Figure 1b). First, the wording of the age question is revised to ask "exact age today." Second, the instructions for filling FOSDIC age circles are revised, and the circles are realigned under the appropriate write-in boxes. Third, we ask for the last two digits of year of birth rather than three digits. The two-digit format is more familiar and should be easier to figure out. Fourth, month of birth was added; this change is not discussed here.

a. Person F's age			b. Month born		c. Year born	
0	0	0	0	0	0	0
1	1	1	1	1	1	1
2	2		2		2	2
3	3		3		3	3

Figure 1b.

Age: Our changes did not significantly reduce item nonresponse for age, which is 6 percent for the 1986 form and 5 percent for the revised form.

However, in computing these item nonresponse rates, we counted items with either FOSDIC or write-in entries as complete. One of the purposes of our revisions was to increase the percentage of persons who filled in both the write-in and FOSDIC portions consistently. Our changes did increase the number of correct responses. As shown in Table 2, for 75 percent of responses on the revised form, all parts were complete and correct--that is, both the write-in and FOSDIC sections were correctly and consistently filled out--compared to 66 percent of responses on the 1986 form.

We also hoped to increase the percentage of respondents who filled the write-in portion correctly, since it must be relied upon if FOSDIC circles are left blank or are indecipherable. The proportion of correct write-in responses is almost ten percentage points higher on the revised form than the 1986 form.

Several kinds of errors occurred equally on the two forms. In two cases statistical differences were observed, and in both cases the revised form did better. Respondents on the revised form were less likely than on the 1986 form to omit a write-in in combination with a correct FOSDIC entry. In addition, these respondents were less likely to make errors in the write-in and omit a FOSDIC response entirely.

To test whether the question was interpreted correctly by respondents, we asked two debriefing questions after completion of the census forms. One question contained a facsimile of the age item and asked, "What do you think is meant by the phrase 'age at last birthday' in Question 5?" (for the 1986 form) and "What do you think is meant by the phrase 'exact age' in Question 8a?" (for the revised form). Results in Table 3 show that a significantly higher percentage of respondents to the revised form correctly understood the question (89 percent vs. 80 percent on the 1986 form).

We also asked a hypothetical question: "Today is John's birthday. He turned 14. How should his mother record his age?" The bottom half of Table 3 shows that 75 percent of respondents to the 1986 form indicated that John's mother should mark his age as 14 (the correct answer), while 87 percent of respondents to the revised form did so.

In sum, our revisions to the age question appear to have reduced some types of errors that could affect FOSDIC-captured data. They did not, however, reduce item nonresponse rates.

Conceptual Problems and Question Wording

Many respondents make errors because questions are vague or use unfamiliar terms, because answer categories do not fit the question, or because questions ask about complex or subjective concepts. Conceptual and wording problems are often hard to discover, because respondents go ahead and answer the question, and are reluctant to reveal their confusion or misunderstanding. The result is that their answers may not mean what we think they mean.

We did some experimentation with question wording on our revised form, in combination with probing for respondents' interpretations of concepts and question wording on the debriefing questionnaire. We present the results of our research on two questions which illustrate different solutions to question wording problems.

Property size: Several wording problems are well-illustrated by a question which at first appears simple: "Is this building on a city or suburban lot, or on a place of less than 1 acre?" One problem is that the question is phrased as an either/or question, suggesting that the respondent should choose one alternative or the other (city or suburban lot, vs. place of less than 1 acre). This is not the case, however; the question is meant to be answered "yes" or "no." In addition, it is double-barrelled, and it is not clear whether people who live on city or suburban lots of more than one acre should answer "yes" or "no." Respondents in our observation sessions misinterpreted the phrase "city or suburban lot" in several ways: some thought the question asked whether the city owns their lot, and others

thought it asked whether they live in the city. Finally, respondents who understand the question still do not know how big an acre is. (City-dwellers, in particular, are ignorant about the size of an acre, which explains why the index of inconsistency for this item is high inside SMSAs, and low outside SMSAs; see DeMaio, 1985.)

This item is a filter question and the data are not actually used, so we dropped it entirely from our experimental questionnaire and replaced it with a skip instruction. We did, however, probe respondents to test their understanding of "acre." First we asked "Do you know what is meant by the term 'acre'?" then we asked respondents to estimate the size of an acre: "Which of the following is closest in size to an acre? a square mile? a football field? a basketball court?" A football field is quite close to the size of an acre; excluding the end zones, it is approximately 5,333 square yards; an acre is 4,840 square yards.

Of those who answered the question, 91 percent claimed to understand the term "acre." (Their understanding of acre did not vary by form type.) However, respondents were overconfident of their knowledge of the size of an acre. Only 46 percent of the respondents who answered the question reported correctly that an acre is closer to the size of a football field. Incorrect responses were more-or-less evenly split between overestimates and underestimates.

Number of rooms: It is sometimes difficult to answer a question, not because the question is unclear, but because it refers to concepts which are inherently difficult. A good example is the question, "How many rooms do you have in your house or apartment?" The question seems so simple and clear that it may be surprising to learn that the index of inconsistency for this item is moderate (45), according to results of the 1970 Content Reinterview Study (U.S. Bureau of the Census, 1975).

Upon examination, the concept of "room" is not as simple as it seems. Respondents commonly misinterpret this question as asking for the number of bedrooms. One respondent in our observation sessions thought this because when realtors talk about the number of rooms in a house, they mean the number of bedrooms. Others saw the list of things not to count and jumped to the conclusion that only bedrooms were to be counted. A second problem is ambiguity about what constitutes a "room" in situations where the space is used as if it were divided even though no walls exist. Some respondents counted combination living-dining rooms as two rooms, even though by the census definition they should be counted as one room unless there is a partition from floor to ceiling. Third, there is ambiguity about whether to count unfinished areas (such as basements) and storage rooms. The census definition of "room" is meant to cover "whole rooms used for living purposes." The 1986 form provides instructions to respondents, along with the question: "How many rooms do you have in your house or apartment? Do NOT count bathrooms, porches, balconies, foyers, halls, or half-rooms." However, we observed

that many people did count rooms which should have been excluded.

We tried to clarify the meaning of "room" by listing examples of both inclusions and exclusions, and we gave an example of an unfinished room, which is not mentioned in the 1986 version. The revised version reads "How many rooms do you have? Count whole rooms, such as living rooms, bedrooms, and full kitchens. Do not count bathrooms, porches, balconies, halls, unfinished basements, and storage areas." We moved the questions on number of bedrooms and bathrooms to immediately follow the question on total rooms, in order to emphasize the distinctions among these questions. We thought that the quality and consistency of the data might be improved by grouping together all questions which require respondents to count rooms.

Our evaluation of the revisions is based on: 1) a comparison of answers to the census items on number of rooms and number of bedrooms; and 2) results of the debriefing.

In most American households, the number of rooms in a house or apartment is larger than the number of bedrooms. Other living arrangements also exist, such as efficiency apartments, in which the number of rooms and the number of bedrooms are equal. But in general, comparison of the responses to these two census items provides some measure of validity.

Table 4 shows that the number of rooms is reported as greater than the number of bedrooms for 92 percent of the revised forms, compared to 83 percent of the 1986 forms.⁴ This result suggests that the information on the revised form is superior. The next comparison is more unexpected: 11 percent of the 1986 forms contain a number of bedrooms larger than the number of rooms. This is an impossible situation, and indicates that one or the other of the responses is inaccurate. In contrast, less than 1 percent of the revised forms fall into this category.

The debriefing questionnaire provides some insight into these findings. In debriefing, respondents were asked to list the rooms they included in their count. Respondents on the revised form, but not on the 1986 form, were explicitly instructed to include living rooms, bedrooms, and full kitchens. Table 5 shows that living rooms and kitchens were counted significantly more often by respondents to the revised form.⁵ This suggests that our explicit instruction to include these rooms improved reporting. However, respondents on both forms were instructed not to count bathrooms. These rooms were also counted by more respondents to the revised form, although the difference between forms is not significant.

The explicit mention in the revised form of rooms to include and to exclude may have stimulated some respondents to count all the rooms that they saw listed and overlook the instruction to count some and not to count others. Whatever the explanation, grouping the items to create a consistent frame of reference for the respondent appears to provide better data. Census reports of number of rooms and number of bedrooms are more consistent, and debriefing results indicate a more complete count of rooms.

Conclusion

Our results demonstrate that a respondent-centered program of research has potential for improving the design of the census long form. Our revised form showed substantial improvements over the 1986 form in several areas, although not all of our initial hypotheses were supported. The results are based on a small nonrepresentative sample of respondents, however, and more research is needed to see whether these results will generalize to the population as a whole. We are currently planning a national mailout/mailback survey to answer this question. In this test we will analyze the effects of our revised wording and sequencing changes separately from our design and layout changes. We hope this test will give us a good basis for making decisions about the design of the census form in 1990 and future decades.

Footnotes

¹This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.

²These figures are calculated from the number of people listed in Question 1 in forms where the item was completed, and the number of person columns filled in forms where Question 1 was blank.

³All differences reported are statistically significant at the .05 level or better using a chi-square test, unless otherwise noted.

⁴All percentages in this section are based on forms in which both the number of rooms and the number of bedrooms items were completed--92 percent of the 1986 forms and 96 percent of the revised forms.

⁵Readers should use caution in interpreting the figures in Table 5, since nonresponse for this debriefing item was high--23 percent for respondents to the 1986 form, and 17 percent for respondents to the revised form.

References

- DeMaio, Theresa. 1985. "Compilation of What We Know About 1980 Census Data." Preliminary Evaluation Results Memorandum No. 105. U.S. Census Bureau.
- _____. 1986a. "Early Results of the Questionnaire Design Project: Observations of Census Employees Filling Out Census Forms." Memorandum. August 13, 1986. U.S. Census Bureau.
- _____. 1986b. "Forms Appraisal of the 1986 Test Census Long Form." Memorandum for Paula Schneider, December 12, 1986. U.S. Census Bureau.
- _____. 1986c. "Forms Appraisal of the 1986 Test Census Long Form." Memorandum for Arthur Young, December 23, 1986. U.S. Census Bureau.
- DeMaio, Theresa, and Elizabeth Martin. 1987. "Documentation of Revisions to 1986 Test Census Long Form (DC-1409, OMB Clearance 0607-0532)." March 16, 1987. U.S. Census Bureau.
- Holt, Mimi, and Judith Lessler. 1987. An Evaluation of the Long Form Questionnaire

for the 1986 Census of Central Los Angeles County. Final Report, Questionnaire Design Project, U.S. Census Bureau Contract No. 50-YABC-6-66032.

Spencer, Gregory. 1985. "1985 Tampa, Florida Test Census - Review of Age Reporting." 1985 Test Census Preliminary Research and Evaluation Memorandum No. 8., U.S. Census Bureau.

_____. 1986. "The Characteristics of Centenarians in the 1980 Census." Paper presented at the April, 1986, meetings of the Population Association of America, San Francisco, California.

U.S. Census Bureau. 1975. Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews. PHC(E)-10.

TABLE 1
Measures of Form Completion

	1986 Form (N = 256)	Rev. Form (N = 259)
Average household size	3.5	3.7
Percent of forms with:		
Question 1 blank	28%	7%**
No sample data	5	2*
Incomplete sample data	16	9**
Number of persons listed:		
In 100% person section	703	743
In sample person section	617	678
Percent of persons for whom we have sample data	88	91

TABLE 2
Summary of Information Related to the Age Item

	1986 Form	Rev. Form
Percent of persons for whom age entries were:		
Correct write-in, correct FOSDIC	62%	71%**
Correct write-in, incorrect FOSDIC	2	4
Correct write-in, no FOSDIC	4	4
Incorrect write-in, correct FOSDIC	5	4
Incorrect write-in, incorrect FOSDIC	5	4
Incorrect write-in, no FOSDIC	12	6**
No write-in, correct FOSDIC	4	1**
No write-in, incorrect FOSDIC	1	1
No write-in, no FOSDIC (item non-response)	6	5
Number of persons	703	743

NOTE: * = Significant at the .05 level
** = Significant at the .01 level

TABLE 3
Responses to Debriefing Questions About Age

	1986 Form	Rev. Form
"What do you think is meant by the phrase 'exact age'/'age at last birthday'?"		
The age you are today	79%	89%**
The age you were before your last birthday	6	2
Don't know	1	-
NA	14	8
Total	100	100
"Today is John's birthday. He turned 14. How should his mother record his age?"		
13	9%	3%
14 (correct answer)	75	87**
15	-	1
NA	16	9
Total	100	100
Number of respondents	256	259

TABLE 4
Comparison of Reported Number of Rooms and Bedrooms

	1986 Form	Rev. Form
Percent of forms on which number of rooms:		
is greater than no. of bedrooms	84%	92%**
is less than no. of bedrooms	11	1**
is equal to no. of bedrooms	5	8
Total	100	100
Number of respondents who answered both items	236	249

TABLE 5
Results of Debriefing Question on Rooms Counted

	1986 Form	Rev. Form
Percent who counted each type of room:		
Kitchen	87%	95%**
Bedrooms	95	96
Living/dining room	89	96**
Bathrooms	10	14
Number of respondents who answered debriefing item	196	214