2010 NRFU Questionnaire Development:
From the 2004 Census Test to the 2008 Dress Rehearsal[1]

Jennifer Hunter Childs

Statistical Research Division
U.S. Census Bureau
Washington, D.C. 20233

---

[1] This report is released to inform interested parties of research and to encourage discussion. Any views expressed on the methodological issues are those of the authors and not necessarily those of the U.S. Census Bureau.

This paper was written prior to the decision not to use the handheld computers for Nonresponse Followup in the 2010 Census. The paper documents research conducted and plans as they were prior to that decision.

Abstract

This paper summarizes the research conducted by SRD between 2004 and 2008 to evaluate and improve the questionnaire planned for use on the handheld computer for the Nonresponse Followup (NRFU) for the 2010 U.S. Census. This paper provides an overview of results of three rounds of English cognitive testing, two rounds of Spanish cognitive testing, two rounds of behavior coding of successive versions of the NRFU instrument (both conducted in both languages), an observational study of the NRFU interview (also conducted in both languages), and two rounds of usability testing with both language instruments. Results from these different pretests led to the development of a revised NRFU questionnaire that was cognitively tested in English. This paper highlights key findings and recommendations from these different pretesting methods. The two largest modifications to the questionnaire based on testing were the creation of a topic-based instrument (as opposed to a person-based one) and the revision of the presentation of residence instructions from a flashcard to a series of questions. The revision to a topic-based sequence allows the interviewer to shorten questions for later people in the household. The need for this change was demonstrated through both cognitive testing and behavior coding. The revision of the residence instruction presentation is aimed at fixing the problem demonstrated through the observational study that interviewers do not often present respondents with the flashcard and through cognitive testing that when presented with it, respondents tend not to read the flashcard. Additionally, the question-answer sequence is expected (based on cognitive test findings) to produce better and more complete results, particularly with respondents with lower levels of literacy in both languages.

Keywords: nonresponse followup; behavior coding; cognitive testing; usability testing; Spanish pretesting.

2010 NRFU Questionnaire Development:
From the 2004 Census Test to the 2008 Dress Rehearsal

As a part of the decennial census operations, the U.S. Census Bureau mails out forms to most housing units in the country. People who do not mail back their census forms are visited by a census interviewer who comes to record their data during a personal interview. This visit is a part of the Nonresponse Followup (NRFU) operation. In planning for the 2010 Census, the Census Bureau moved from collecting NRFU data using a paper data collection instrument to collection via handheld computer (HHC).[2] This change in mode of data collection necessitated some changes in the question text in order to optimize the interviewer and respondent interactions. During 2004 and 2006, the Census Bureau tested this new data collection in field tests.

The Census Bureau migrated from a paper-and-pencil form to an HHC for several reasons, including the reduction of paper, the ability to encourage standardized interviewing procedures by automating complex survey paths and "fills," and the reduction of cost (for historical context of moving to a Computer Assisted Personal Interview instrument, see Nicholls and de Leeuw, 1996; for more immediate reasons see Waite, 2007). To our knowledge, the testing cycle of the NRFU HHC is one of the first attempts by a U.S. government agency to collect actual survey data using an HHC (see Bosley, et al. 1998; Couper, 2002). However, other researchers in academia, the private industry, and other countries have been using HHCs to gather data for the last 10 years (e.g., de Heer, 1991; Nusser, et al. 1996; Greene, 2001; Gravlee, 2002; Ice, 2004; Cameron, 2005; Johnson & Horne, 2005). Nusser (1999) notes that the development of mobile-Computer Assisted Personal Interview (m-CAPI) devices for human population surveys has trailed far behind the use of such devices for the gathering of environmental data. Hence, as far as we know, this is one of the first evaluations of question wording presented on an HHC to household respondents.

The Census Bureau progressed through a series of field and other qualitative pretests that culminated in an m-CAPI data collection instrument that takes advantage of the benefits offered by the technology while maintaining comparable data quality to the self-administered data collection. This paper documents the progress of the NRFU development.

NRFU Background

The NRFU instrument collects very basic data on the housing unit (e.g., whether the unit is occupied or not, whether the unit is owned or rented) as well as some basic demographic data about each person who lives in the household (e.g., names, ages, race). The demographic data could be (and during the testing cycle, it was) collected using two different strategies. The first is a person-based approach, which consists of a series of questions that are asked in their entirety about the first person with whom the interviewer speaks. Then the same series is administered again about the next person, and so on (e.g., sex, age, date of birth, and race data are gathered about Person 1, and then data on the entire series are gathered about Person 2). The alternative method of administration is topic-based, in which data regarding a single topic are gathered for

---

[2] The HHC used for data collection during these tests was an electronic device operated by using a stylus; such devices are also known generally as personal digital assistants (PDAs).

everyone in the household before moving on to the next topic in the survey (e.g., race is gathered for everyone in the household, then age is gathered for everyone in the household).

The 2004 NRFU instrument was the first HHC developed for census data collection and was developed initially using the 2004 self-administered paper form. It was person-based and had questions that mimicked the self-administered form very closely. The English and Spanish versions of the 2004 instrument were evaluated using the behavior coding method described more fully below.[3]

During this decade, the NRFU operation has implemented flashcards. These are visual aids used by interviewers to assist respondents when the interviewer needs to present more information than can be stored in working memory. Flashcards were used to present the residence instructions, relationship, Hispanic origin and race response categories. See Appendix A for an example of the flashcards used in the 2006 test.

Based in part on results from the 2004 behavior coding, cognitive testing with the mailout form, and input from the Bureau's survey methodologists, the NRFU questions were modified between the 2004 and 2006 tests. Among the changes was a shift from person- to topic-based format. The revised 2006 NRFU instrument was pretested using behavior coding as well as cognitive and usability testing, in English and in Spanish.

Based on the two rounds of behavior coding, and two rounds each of English and Spanish cognitive testing with the NRFU instrument, two rounds of usability testing, and an evaluation study of the residence rule and instructions, a revised NRFU script was developed.[4] The revised script was tested using a single round of English cognitive testing (a translation was not yet available). Based on findings from that study, the revised questionnaire was prepared for the 2008 Dress Rehearsal.

The goal of this paper is to provide a big picture overview of the evolution of the NRFU instrument from the 2004 script, based heavily on the paper self-administered questionnaire, to the 2008 Dress Rehearsal script which takes advantage of the technology offered by the HHC.

Where appropriate, I reference the Mode Consistency Guidelines (see Martin et al., 2007), which were developed to ensure consistency between different modes of data collection for the 2010 Census.[5] In conjunction with experts in the field of survey methodology, the Census Bureau developed guidelines for questionnaire development that aim at maintaining consistent data quality across multimode surveys. The impetus for this effort was the multiple modes of data collection that will be used for the 2010 Census and the American Community Survey, including self-administered questionnaires and telephone and personal visit interviews conducted using both paper and electronic instruments. This effort resulted in the creation of 30 guidelines that

---

[3] There is a Spanish version of the NRFU instrument and flashcards available for interviewers who are able to conduct interviews in Spanish.

[4] Each of these methods will be described in more detail in the following sections.

[5] Though the Mode Consistency Guidelines do not represent a pretest conducted in the development cycle of the NRFU, they did carry great weight in the decision making process that led to many of these changes. For that reason, they will be described here to the extent they are relevant.

introduce the principle of Universal Presentation and apply it to several major aspects of instrument design including question wording and instructions, flashcards, and edit messages. There are apparent exceptions to the principle of Universal Presentation involving situations in which a change in the question wording, order, or instructions is either essential for operational reasons, or to better preserve the question than would asking the question identically in different modes. See Martin et al. (2007) for more details.

These guidelines referred to above were developed concurrently with many of the research activities described below. I will refer to these guidelines as needed to supplement findings from the research described here.

Methods

In order to see the full picture of the development of the m-CAPI survey from a paper-administered form, this paper briefly describes each study, as well as the findings that were pertinent to making the changes that resulted in the 2008 Dress Rehearsal questionnaire. For more information on any of the studies contained here, see the corresponding study reports.

Behavior Coding

Behavior coding is the systematic coding of interviewer and respondent interactions (Cannell, Fowler, & Marquis, 1968). It identifies flawed questions by revealing administration and response issues. Problems are detected by looking at rates of undesirable behavior that exceed 15 percent (Oksenberg, Cannell, & Kalton, 1991; Fowler, 1992). Census telephone interviewers are trained in project-specific techniques to apply behavior codes while listening to audiotaped interviews. The coders are trained to take detailed notes whenever a non-ideal interaction occurs. Qualitative analysis of coders' notes allows us to see exactly where problems occur and allows us to hypothesize how these problems might be solved. Looking at behavior coding data, we focus on three major behaviors: interviewer behavior, a respondent's first response behavior, and whether or not the respondent interrupted the interviewer during the reading of the question (we call this a break-in).

Behavior coding was conducted during both the 2004 and 2006 Census Tests. In 2004, we gathered 220 audio-taped interviews (119 English, 72 Spanish, and 29 mixed English and Spanish). In 2006, we gathered only 72 usable audiotapes (due to a problem with the contractor who was hired to do the collection). The majority of these were English (54), but analysis was also conducted on the 18 usable Spanish tapes. These results are documented fully in Hunter and Landreth (2005) and Childs, Landreth, Goerman, Norris and Dajani (2007), respectively.

Cognitive Testing

Cognitive testing is a method by which participants are administered a survey, usually in a lab setting, and are asked semi-scripted concurrent or retrospective probes about their thought processes while answering the questions. Results from cognitive testing show us where respondents in a production survey may have difficulties or answer incorrectly. See Willis (2005) for a detailed explanation of cognitive testing as a pretest method.

Cognitive testing was conducted using the 2006 NRFU script. Two rounds of English testing were conducted. The first 14 interviews were conducted using a paper script and are documented in Hunter (2005). The second round consisted of 16 interviews, conducted using the 2006 NRFU instrument as it was programmed in the HHC. Results from the second round are documented in Childs, Gerber, Carter and Beck (2006).

The Spanish script of the 2006 instrument was also tested in two rounds. Both rounds of 15 interviews were conducted by contractors and used paper scripts of the instrument. The first round is documented in Beck (2006) and the second in Jones and Childs (2006).

The final round of cognitive testing, with the revised, recommended NRFU script was conducted in English only. It consisted of 28 respondents, all living in complex living situations (to provide a good test of the revised residence rules questions). This study is documented in Childs, Carter, Norris, Hanaoka, and Schwede (2007).

Observational Study

Development Associates was hired under contract with the Census Bureau to conduct an observational study of the 2006 Census Test NRFU in Travis County, TX. The goal of the study was to understand flashcard use and other non-verbal behaviors that would not be available using the audiotapes used for behavior coding. The contractor observed 99 eligible interviews, 65 in English and 34 in Spanish. These findings are documented in Rappaport, Davis and Allen (2006). A subset of these interviews was taped for the behavior coding described above.

Usability Testing

The goal of usability testing is to improve the usability of the product so that "the *people who use the product* can do so *quickly and easily*" (p. 4, Dumas and Reddish, 1999). In usability testing, the participant plays the part of an interviewer, given a limited amount of training on the instrument. The "respondents" are confederates played by the researchers using prearranged respondent scripts. The goal of a usable instrument is to be intuitive enough for someone with limited training to be able to navigate without many problems.

Two rounds of usability testing were conducted with two early cuts of the 2006 NRFU instruments in July and August of 2005. The first round of usability testing had 6 participants (4 English-speakers and 2 Spanish-speakers) and is documented in Olmsted, Hourcade and Abdalla (2005). The second round had 5 participants (4 English-speakers and 1 Spanish-speaker) and is documented in Olmsted and Hourcade (2005).

Evaluation Report

One of the official evaluations of the 2006 Census Test examined the performance of the residence rule and residence instructions (Heimel, King and Sheppard, 2007). This evaluation consulted the observation and behavior coding studies described here, as well as structured field observation reports, trip reports and interviewer debriefings. See the full report for details on

their methods and full results. The study is mentioned here because it corroborates many of the findings from the studies described in this report.

General Findings

Behavior Coding

In the 2004 Census Test, in which the NRFU was person-based, we found that on average, interviewers only asked questions as intended 36 percent of the time (Hunter & Landreth, 2005). Every question was asked correctly more often the first time than later times (first time - good interviewer behavior averaged 47 percent; later times - good interviewer behavior averaged 32 percent). Though we hope interviewers always read questions as worded, this finding is not too surprising given that the same questions are repeated for each person in the household. We hypothesized that the poorer interviewer behavior for later persons might be due to interviewers compensating for information they thought they already had, either the respondent provided the information explicitly (e.g., the respondent said "we're all white"), or the interviewer assumed it based on answers given for other household members.

Although there were no significant effects of person number on respondent interruptions, we noted that for 6 of the 7 questions the trend was that respondent interruptions were more frequent for later administrations than for the first administration (Hunter & Landreth, 2005). This further demonstrates that repeating the same questions over and over for each household member can become burdensome. These findings, along with literature from survey methodology (see Moore & Moyer, 2002; Colosi, 2001), led to the change from a person-based administration in 2004 to a topic-based one in 2006.

In the 2006 Census Test, the topic-based instrument still fared somewhat worse than we had hoped. Ideal interviewer behavior only occurred, on average over questions, in 40 percent of administrations (Childs, Landreth, et al., 2007). The scripted question text expected interviewers to read the question in full for each person, despite the topic-based format which could have allowed shortened questions after the first full reading of the question (e.g., What is Mary's age? How about Susan?). Unsurprisingly, administration was better for the first time than for later times, meaning that interviewers were more likely to read a question as worded the first time. Average good interviewer behavior was 50 percent for the first time, and 32 percent for later administrations. Most major changes for later administrations involved "how about ___?" wording. Based on these findings, we recommended implementing "full" topic-based administration by allowing these shortened questions for later people when the context has not been interrupted. This recommendation was adopted for 2008.

We also noticed differences in good interviewer behavior between the English and Spanish versions of both the 2004 and 2006 instruments (Hunter & Landreth, 2005; Childs, Landreth, et al., 2007). In both instruments, for each question examined, the English version had higher (or equal in one case) rates of good interviewer behavior than the Spanish version of the question. This means that interviewers were better able to read the English questions as intended than the Spanish ones. We attribute these differences to complex English wording which becomes more complex through translation, inexact translations, and errors in the Spanish translation instrument

that was fielded. See the individual reports for more information on this topic (Hunter & Landreth, 2005; Childs, Landreth, et al., 2007).

Cognitive Testing

Most findings from the cognitive testing were question-specific. A few general findings involved interviewer burden. In many cases, it was obvious from cognitive testing that in a topic-based administration, shortened questions for later people would decrease interviewer and respondent burden substantially without disturbing data quality. This recommendation was implemented in the 2008 NRFU specification for the most part, with the exception of scripting the reading the race question in its entirety twice due to the recommendation of subject matter experts.

Several of the questions (e.g., tenure and overcount) in the 2006 NRFU were scripted too long for oral presentation. These were tailored in the 2008 NRFU script to improve interviewer adherence to the script (see Question-by-Question analysis for more detail).

Finally, regarding the questions that needed to refer to the housing unit, asking interviewers to choose the appropriate term among "house/apartment/mobile home" was too burdensome. Because the interviewer uses an automated instrument, the answer could easily be an automatic fill. This was a general recommendation that was implemented in the 2008 NRFU specification.

Observational Study

The observational study provided our most comprehensive examination of interviewer behavior with flashcards to date. Out of 99 interviews observed, the Residence Rules flashcard was presented in 25 percent of cases, the Relationship flashcard in 28 percent of cases and the Ancestry flashcard in 37 percent (Rappaport, et al., 2006). Interestingly, in 45 percent of observed cases at least one flashcard was used – which indicates that interviewers pick and choose which flashcard should be used in an interview. This did differ by language. In English, the cards were used at rates of 28 percent, 25 percent, and 38 percent, respectively. In Spanish, the rates were 17 percent, 33 percent, and 33 percent, respectively. Interestingly, the residence rules card was used somewhat less in Spanish interviews, despite the fact that non-English speaking immigrants often live in mobile, complex households (Goerman, 2005).

In response to these findings, similar findings from an evaluation study (Heimel, King & Sheppard, 2007), and the findings described in more detail in the Question-by-Question analysis, the residence rules flashcard was dropped for the 2008 NRFU.

However, there was still a need for flashcards for the Relationship, Hispanic origin and Race questions for 2008. Because of the documented difficulty interviewers have with using flashcards (and because we know from interviewer debriefings that they often chose not to use them), we revised the format of the flashcards. It was noted during observations of the field tests that interviewers did provide respondents with our legally required "confidentiality notice." Because we observed interviewers handing respondents that notice, and not using the flashcards, we thought we could take advantage of the former, and provide respondents with a single "information sheet" that contains the confidentiality notice, as well as "lists" for the relationship,

Hispanic origin, and Race questions. This new idea will be tested during the 2008 Dress Rehearsal.

Usability Testing

The usability studies provided corroborating information to the cognitive test findings. Olmsted et al. (2005) noted the repetitiveness in reading topic-based questions over and over for each household member. Olmsted and Hourcade (2005) mentioned the difficulty of working with flashcards when their use is not scripted in the instrument itself. Those authors also recommended automatically filling "house/apartment/mobile home" when possible.

Evaluation Study

The evaluation study provided more evidence that flashcards are seldom used. One of the studies reported by Heimel and colleagues (2007) found that the residence rules flashcard was shown in 41 percent of interviews. Heimel's study also illuminated reasons why flashcards were seldom used by interviewers – some interviewers thought they took too much time to use; others thought they bothered the respondents, and might cause them to lose the interview; and others stated it was too much to hold.

<div align="center">

Question-specific Findings

</div>

Coverage

This section will discuss the questions from the NRFU instrument that are targeted to count each person once, only once, and in the right place. Determining where to count a person on census day is typically a straightforward task, since most people have only one home, and they stay at that home most of the time. However, for individuals who stay at multiple places, whether they have an official second home or they frequently stay informally with friends or family, where to count them in the census becomes more complicated. For the most part, the Census Bureau employs a *de jure* rule to count people at their usual residence, that is, where they are usually living and sleeping, on census day[6]. Past research has shown that respondents' understanding of where each person *lives* often differs from the Census Bureau's rule of usual residence (Gerber, 1994). Because of these complexities, the question sequence that allows us to determine residence is of particular interest to the Census Bureau.

Determining the Path

Prior to starting the interview proper, a question must be asked to figure out if the interviewer is attempting an interview with a census day resident or a household resident who moved in after census day. In 2004 and 2006, the original question read like this:

---

[6] There are a few important exceptions to this rule. The Census Bureau enumerates people who stay in group quarters (i.e., places that house groups of people like college dormitories, nursing homes and jails) in the place where they stayed on census day. This particular rule is *de facto* rather than *de jure*. See the report from the National Research Council (2006) on the complexities of applying this rule.

Did you or anyone in your household live at [ADDRESS] on April 1, [YEAR]?

Behavior coding identified errors by the interviewer that included leaving out the date, replacing the address with "here" and occasionally providing a spontaneous explanation (Childs, Landreth, et al., 2007). These are all indications of problems with the question suggesting that the question is too long or complex. In cognitive testing, it was noted that sometimes respondents did not list themselves on the roster (Hunter, 2005). This was attributed to the fact that the respondent had already reported (via this question and one earlier in the instrument) that he or she lived there on April 1st. The roster question, then, was interpreted as asking for "who else lived there on April 1st?". This will play a role later, as well, when the roster is discussed.

The recommendation tested during the final round of cognitive testing separated this question into two parts. The script first asks if the person lived there on census day, and then, if not, whether anyone else in the household lived there on census day. This accomplishes two things. First, it allows the computer to set a flag as to whether or not the respondent should be on the roster. Second, it simplifies the task by asking the straightforward question of whether the respondent was there on census day, and then asking the more complicated question as to whether anyone living there now was also living there on census day. The questions tested are as follows:

1. **Did you live at <Address> on April 1, 2006?**
   Yes – Go to 3 (lets us know the respondent is a household member)
   No - Go to 2

2. **Is there anyone living <here/ at Address> now who also lived here on April 1, 2006?**
   Yes – (continue, but don't assume the respondent will be on roster)
   No – *Proxy interview*

These questions tested well in our cognitive testing with nonproxy situations (Childs, Carter, et al., 2007). Based on observations of the 2006 Census Test Census Coverage Measurement Person Interview, we recommended adding "stay" to this question to include respondents who do not consider themselves permanent residents of the household, but who stay there most of the time (see Nichols et al., 2007). The final recommended wording for these questions for the 2008 NRFU was as follows:[7]

1. **Did you live *or stay* at <Address> on April 1, 2006?**
   Yes – Go to 3
   No - Go to 2

2. **Is there anyone living *or staying* <here/ at Address> now who also *stayed* here on April 1, 2006?**
   Yes
   No – *Proxy interview*

---

[7] Appendix B shows a simplified personal-visit NRFU script as it is specified to appear for 2010. This may be used for context while reading about the question-by-question findings and to see the "big picture" of the revisions.

Usual Residence

The usual residence question is aimed at determining whether the housing unit is occupied as someone's usual residence, or as a vacation, seasonal or temporary home. If it is only a temporary residence, the Census Bureau will not enumerate people there (as people should be enumerated where they usually live and sleep, per our residence rule). In 2004, this question was scripted:

> **Is this (house/apartment/mobile home) the usual residence of someone in your household, or is it a vacation home, seasonal home, or held for occasional use?**

During the 2004 behavior coding, this question was frequently reworded by the interviewers, who in many cases shortened it to ask if the place was the usual residence or not (leaving out the other alternatives; see Hunter & Landreth, 2005). Based on previous research (Gerber, 1994), we know that respondents have difficulty with the term "usual residence." Thus, shortening the question in this way may leave respondents not knowing how to answer. As evidence of this, the behavior coding showed high rates of inadequate answers from respondents. Instead of responding to the question as it was asked, respondents often said it was their "regular" or "permanent" home, indicating that they had not picked up on the use of the term "usual residence." In a few cases, respondents even answered inappropriately by saying that they rent the home, or just answering "no."

We had the opportunity to cognitively test the 2004 wording during the first round of the cognitive testing. Those findings supported behavior coding findings in that most respondents did not remember the term "usual residence" by the end of the question (Hunter, 2005). Respondents instead answered with words like "regular" or "permanent." Additionally, several Spanish-speaking respondents needed to hear the question more than once to understand its intent and other Spanish-speakers answered "no" – incorrectly thinking this question only dealt with temporary residences (Beck, 2006).

In 2006, the question was revised to remove the words "usual residence" from the question text. Inadvertently, however, "usual residence" remained in the response options, and we believe this impacted interviewers' behavior. Additionally, the most common option "usually live here" was moved to the end of the question to take advantage of recency effects in the aurally administered question (see Tourangeau, Rips and Rasinski, 2000). The 2006 question read as follows:

> **Is this (house/apartment/mobile home) a vacation home, seasonal residence, held for occasional use, or does someone in this household usually live here?**
> [] Vacation home, seasonal residence, held for occasional use
> [] Usual residence

Comparisons of behavior coding statistics from before and after this change suggest that this resulted in an increase in exact reading (32 percent to 51 percent, though significance testing is not meaningful across these two very different test situations; Hunter & Landreth, 2005 and Childs, Landreth, et al., 2007). However, the rate of exact reading was still very low with the revised question. One common problem was for the interviewer to pick up the term "usual

residence" from the response category and probe with that rather than asking the scripted question. It was noted that some interviewers emphasized or repeated the "usually live here" option – to try to encourage respondents to answer that way (as it was most likely the correct response). Interviewers still often left out phrases of the question. Respondents still had difficulty providing responses in the terms of the question.

Cognitive testing with the 2006 question also showed it improved over the 2004 version (Childs, et al., 2006). Most respondents answered with the term "usually live here," matching the question text. Placing the most likely answer at the end did seem to help respondents remember it, and seemed to help respondents understand what the point of the question was.

Several cognitive test respondents in each round did not know what was meant by the phrase "held for occasional use" (Hunter, 2005; Childs, Gerber, et al., 2006) Additionally, the Spanish cognitive testing yielded the finding that the question was still too long and caused problems with respondents not being able to parse the response categories (Jones & Childs, 2006). Because of the rarity of interviewing at a place "held for occasional use" and the persistent problem with the question length, we recommended dropping the option of "held for occasional use" from the question text and offering it only as a part of the response categories. We also shortened the question slightly. The following wording was tested in the final round of cognitive testing:

> **Is this (house/apartment/mobile home) a vacation or seasonal home or does someone usually live here?**
> __ Vacation, seasonal, held for occasional use
> __ Someone usually lives here

This revision performed fairly well with most respondents understanding and answering as intended (Childs, Carter, et al., 2007). There were still a few problems, however. A few respondents did not know if the question applied to just themselves or to the whole household (in the case that the respondent is the more tenuously attached person). A couple other respondents tried to report seniors or public housing here. Because the question has become much shorter, we recommended rearranging the question to put the most likely option first (now that the question should not tax working memory). Here was the final recommendation:

> **Does someone usually live at this <FILL house/apartment/mobile home> or is this a vacation or seasonal home?**
> __ Someone usually lives here
> __ Vacation, seasonal, held for occasional use

However, this recommendation was not taken and this version appears in the 2008 NRFU specification:

> **Is this <FILL house/apartment/mobile home> a vacation or seasonal home** [pause] **or does someone usually live here?**

Residence Rules

There are particular situations where we know from past research that coverage errors are made with respect to who to list on a census questionnaire. These errors include forgetting to list children, unrelated people and distant relatives as well as erroneously including people who live elsewhere like college students, commuter workers and people who are in a group quarters on census day. To mediate these errors, in 2004 and 2006, the residence instructions were presented to respondents by showing them a flashcard which displayed the instructions, essentially identically to those on the mailout form.[8]

Using a flashcard to convey the residence instructions was demonstrated to be problematic for a number of reasons: 1) in interviewer debriefings, a very high percentage of interviewers readily admitted not using the flashcard (Heimel et al., 2007); 2) observations of the 2006 Census Test showed that only a quarter of interviewers (while being observed) used the flashcard (Rappaport et al., 2006); 3) cognitive and usability testing showed difficulty with the administration of the flashcard because there was no introduction and respondents felt like there was not time to read the card because the interviewer immediately asked them a question (Hunter, 2005; Childs, Gerber, et al, 2006; and Olmsted & Hourcade, 2005); and 4) we know that at least a portion of the hard-to-enumerate population is not very literate, making the flashcard irrelevant to them (National Assessment of Adult Literacy, 2006). Presenting the residence rules is critical to getting accurate coverage of the population. If a large portion of our respondents are not exposed to the residence rules, we will almost certainly have increased errors in our census.

Based on these findings, and the findings presented below, a new strategy for presenting residence instructions was developed. The new strategy involves a series of questions that provide the same information as the residence rules and coverage questions on the mailout form, but adapt the questions to be suited for an interviewer-administered format. By combining the residence rules presentation and the coverage questions, the redundancy that became obvious in the interviewer-administered script is eliminated. Table 1 below presents an overview of the 2004/2006 series of questions and the revised 2008/2010 series of questions for reference in the discussion that follows.

---

[8] See the 2006 Who to Count flashcard in Appendix A.

Residence Instructions/ Coverage Questions	Previous 2004-2006 Format	Revised 2008/2010 Format
Residence Rules	Displayed on Flashcard	No flashcard - Presents residence rules through series of questions
Count	Asked directly	Calculated from people listed
Roster	After reading the flashcard and giving count, list all people starting with the owner/renter	Start with respondent in non-proxy situation– any name in a proxy situation – List all household members
Reference Person (owner or renter)	Embedded in the Roster question	Asks a separate question to determine reference person
Undercount	Reiterates the same types of people from flashcard and asks if we missed any of those people, then adds them to roster	Eliminates redundancy – Probes only once for people we know are missed in the census and adds them to the roster
Overcount	Reiterates the same types of situations from flashcard and asks if each person sometimes lived somewhere else	Eliminates redundancy – Asks household level questions about types of places from former flashcard and asks individual level questions about other places where people can be duplicated

Table 1. Structure of residence instruction presentation and coverage questions from the 2004/2006 approach to the 2008/2010 approach. [9]

Population Count

The questions that surround the population count have also caused problems for interviewers and respondents. In 2004, the question was read exactly as it should have been only 57 percent of the time (Hunter & Landreth, 2005). In 2006, it was read exactly as worded 52 percent of the time (Childs, Landreth, et al., 2007). In 2004, the housing count question was shorter, not conveying any of the content from the residence rules flashcard. In 2006, two "centralizing bullets" were added to the question text to convey the basic concept of our rules (acknowledging that the flashcard was not often used). However, the lengthened question text led to a slightly higher rate of changes.

2004
How many people were living or staying in this (house/ apartment / mobile home) on April 1, 2004?

---

[9] The question text in the table does not reflect the actual question text in the instrument, but is an abbreviated format for illustrative purposes.

2006
> **The census must count every person living in the United States on April 1, 2006. We want to count people where they usually live and sleep. For people with more than one place to live, this is the place where they sleep most of the time. How many people were living or staying in this [house/apartment/mobile home] on April 1, 2006?**

Cognitive testing of the 2006 version also provided evidence that the question was now too complex (Hunter, 2005). Several respondents answered incorrectly – one leaving himself out of the count, another answering only "one" because she thought it meant other than the people who normally live there.

The recommendation from all of these findings was to use a series of shorter questions to convey residence rules in an interviewer-administered instrument. The revised design does not collect a count of the people living in the unit prior to gathering names. Rather than focusing on counting people at the beginning, the new design uses a question-answer sequence to build a roster, asking, as probes, the topics from the residence rules box on the census form. The housing count is then calculated by the number of people for whom data are gathered.

The revised script that was tested in the final round of cognitive testing will be presented along with findings from the next question.

Gathering Roster (Names)

The question used to gather household resident names in 2004 and 2006 actually had a dual purpose. Mimicking the mailout form, the NRFU asked the respondent to list all the people living or staying at the housing unit on census day starting with the householder (the owner or renter of the unit). It is necessary to identify the householder because household relationships are reckoned to the householder. In census data analysis, statistics on family or nonfamily households are derived from this relationship data.

2004 Question
> **What is the name of each person who lived or stayed at this residence on April 1, 2004? Start with the name of one person who owned or rented this [house/apartment/mobile home] on April 1, 2004.**

2006 Question
> **What is the name of each person who lived or stayed at this [house/apartment/mobile home] on April 1, 2006? Start with the name of one person who owned or rented this [house/apartment/mobile home].**

Findings for this question were among the most surprising. In both field tests, in over half of the behavior coding cases, interviewers read the question with a major change (61 percent and 64 percent in 2004 and 2006, respectively; Hunter & Landreth, 2005 and Childs, Landreth, et al., 2007). In a startling number of cases in both tests, the interviewer asked the respondent to start

with him or herself or did not provide respondents with instructions on whom to start with. This could result in household relationships being reckoned to the wrong person (e.g., to the child or the unrelated boarder). We have no data on whether or not respondents listed the appropriate person first, but the fact that the interviewers had such difficulty with this complex instruction indicates that it is not effective. The usability study pointed out this difficulty as well (Olmsted & Hourcade, 2005).

Additionally, in cognitive testing, we noted that respondents sometimes failed to include themselves on the roster (Hunter, 2005). We attribute this to the complexity of the instruction and the context effect of having already stated that the person lived at this place on April 1st (at the question to determine proxy or nonproxy path).

For these reasons, in the final round of cognitive testing, we tested gathering the respondent's name first in nonproxy interviews, and then asking a separate question to identify the householder. The question that identifies whether the interview will be a proxy or nonproxy interview tells us whether or not the respondent was living at the unit on census day. Based on that flag, we can ask the respondent to start with him or herself, eliminating the risk of the respondent incorrectly leaving him or herself off the roster. This series performed well during cognitive testing (Childs, Carter, et al., 2007).

The final improvement made to this question was to probe for middle initial. Based on cognitive testing in both English and Spanish, we know that respondents usually provide only first and last names to the NRFU interview (paternal name in Spanish; Beck, 2006; Childs, Gerber, et al., 2006; Hunter, 2005; Jones & Childs, 2006). This is consistent with NRFU field test data as well (Norris, 2005). In the last round of cognitive testing, we tested prompting the respondent for "full name," but still did not see an increase in reports of middle initial. Because middle initial is a very good piece of information to use to identify people mistakenly listed twice in the census, and because it is relatively easy to see on a self-administered form that middle initial is requested, we recommended probing for middle initial. This was implemented in the 2008 NRFU specification.

The final recommended wording, which appears in the 2008 NRFU script incorporates a basic description of our residence rules, starting with the respondent in nonproxy cases, and probing for middle initial. It reads as follows:

> **We need to list people living or staying (here/ at this house/apartment/mobile home) on April 1st, 2006. We want to list people where they usually live and sleep. For example, college students and Armed Forces personnel should be listed where they sleep most of the time.**
>
> If nonproxy:
> **Let's start with you, what is your first name? Middle initial? Last name? What is the first name of the next person who was living and sleeping here on April 1st? Middle initial? Last name? Anyone else?**

If proxy:
**What is the first name of a person who was living and sleeping here on April 1st?
Middle initial? Last name? Anyone else?**

According to the Mode Consistency Guidelines (Martin, et al., 2007), this is an acceptable adaptation to the questions on the mailout form because of the preponderance of evidence that the residence rules were not being attended to by respondents in the way they were previously presented. Because respondents were not presented with the same information in the NRFU as they were in the mailout form, modifications needed to be made to the NRFU instrument to ensure all information was conveyed. This resulted in changes to the question sequence for NRFU. Additionally, it was agreed that it is not necessary to collect a housing count from the respondent when one can be calculated after asking the appropriate roster questions.

Undercount Question

The next question asked in the 2004 and 2006 NRFU is the question aimed at identifying possibly omitted people (or census undercount). In 2006 (as is planned for 2008 and 2010), "yes" responses to this question are followed up by gathering names and demographic data for these people. This question also flags people for followup to determine residence status during the Coverage Followup (CFU) operation. The CFU interview is a more in-depth assessment of residence that is used to clarify residence status for households who have indicated through their initial census response that there may be a complex living situation in their household. There are a number of ways that a household can be flagged for CFU, but the ones I will discuss are those that result from the two coverage questions that are a part of most decennial data collection instruments. The first of these coverage questions is aimed at finding out if there are additional people who might be household members according to the census residence rule. This question highlights people who are often left off the census questionnaire. If the respondent answers yes to this question, a flag is set so that an interviewer will return to clarify the residence situation through the CFU questionnaire.

2004[10]

Besides the name[s] you gave me earlier, were there other people who lived or stayed at this place part of the time but were not permanent residents? For example, live-in employees or children in joint custody?
[] Yes
[] No

2006

We do not want to miss any people who might have been staying here on April 1, 2006. Were there any additional people staying here that you did not include, for example:

---

[10] Based on research outside of the scope of that described here, the question changed structure between 2004 and 2006. Since the 2006 version is the approach that we are moving forward with, it will be discussed in more detail here. The 2004 version is shown for reference.

- **Children, such as newborn babies or foster children?**
- **Relatives, such as adult children, cousins, or in-laws?**
- **Nonrelatives, such as roommates or live-in baby sitters?**
- **People staying here temporarily?**

In the 2006 wording, interviewers had difficulty with the length of the question. Interviewers most often left out some, or parts of some, of the response options (Childs, Landreth, et al., 2007). The most interesting finding from the 2006 behavior coding of this question was that in 6 of 72 cases, the respondent mentioned someone who could have been omitted and the interviewer did not list them without probing enough to find out whether or not they should have been listed. Since this question is aimed at finding people who were omitted, interviewers that ignore people mentioned here are making a critical error.

Problems identified during cognitive testing of the 2006 version included respondents reporting people who clearly had a permanent home elsewhere, and also reporting "yes" concerning people who had already been rostered (Hunter, 2005; Childs, Gerber, et al., 2006). The latter is only a big problem if interviewers enter the same names again (causing within-household duplication). The former could cause between-household duplication if CFU does not correctly remove these people from the roster.

This question also elicited some irritation by respondents in cognitive testing and some suspicion that these questions were aimed at finding people who were being untruthful (Beck, 2006; Childs, Gerber, et al., 2006). We think this is because the residence rules flashcard, presented in cognitive testing, provides exactly the same information. If respondents attended to that, we are, indeed, asking them if they followed the directions or not.

The recommendation for this question was to blend it with the presentation of the residence rules, since there is considerable overlap, and to try to shorten the probes as was done for the 2005 CFU questionnaire. This is the question tested in the last round of cognitive testing:

a. We do not want to miss any people who might have been staying here around April 1st. Were there any additional people that you didn't mention, for example:

Babies?
Foster children?
Any other relatives?
Roommates?
      Yes – **What is that person's name? Anyone else?**
      No **-** Continue

b. How about anyone staying here on April 1st who had no other permanent place to live?
      Yes – **What is that person's name? Anyone else?**
      No **-** Continue

c. Anyone who you don't think of as part of your household, but stays here most of the time?

> Yes – **What is that person's name? Anyone else?**
> No **-** Continue

The probes for the initial undercount question are shorter than those tested previously, focusing on the categories mentioned in the residence rules as types of people often forgotten. Questions "b" and "c" are pulled out as separate questions, since their goal is somewhat different (i.e., not people who were forgotten, but people the respondent might not have thought he or she should include). These were tested in the final round of cognitive testing (Childs, Carter, et al., 2007). During cognitive testing, in response to Question "a" respondents mentioned exactly the type of people who should be followed up during CFU. Of seven people mentioned in cognitive testing, three should definitely have been added to the household roster, two were tenuously attached and at risk of being missed by the census all together and two should likely have been counted elsewhere. As a flag for CFU, we think this will be very successful.

Question "b" also worked as intended in cognitive testing, with one respondent actually adding a family who was temporarily experiencing homelessness who had been staying with her around census day. Other respondents understood the gist of the question and could accurately describe the type of people we were looking for (e.g., "homeless people," people who were between places to live or who were "down on their luck"). This question represents part of the residence rules and is recommended to function as a direct add to the roster for NRFU (not as a flag for CFU). By definition, the Census Bureau wants to count people who had no other permanent place to live on census day at the unit where they were on census day.

Question "c" did not fare well in cognitive testing. No additional people were added because of this probe and respondents could not tell how this question was different from the other two. Thus, it was dropped from the sequence in the final recommendations. Questions "a" and "b" above replace the 2006 undercount question for the 2008 NRFU. One additional minor change was the addition of the category "any other nonrelatives" to more closely match the categories on the mailout form. The 2008 question wording is as follows:

> **We do not want to miss any people who might have been staying here on April 1st.**
> **Were there any additional people that you didn't mention, for example:**
>
> **Babies?**
> **Foster children?**
> **Any other relatives?**
> **Roommates?**
> **Any other nonrelatives?**
> **How about anyone staying here on April 1st who had no other permanent place to live?**
>
> Yes – **What is that person's first name? Middle initial? Last name? Anyone else?**
> No **-** Continue

Addressing mode consistency again, evidence shows that the lengthy undercount question was not being administered properly in the NRFU. In order to counteract this, the 2008 NRFU undercount wording was simplified so that we can have more control over what the interviewer says, given that we know interviewers often will not read lengthy questions. This question also incorporates information presented in the residence rules box on the mailout form, thus improving the delivery of that information.

Overcount Question

Rounding out the residence rules and coverage questions, I will present results on the overcount question now, despite that it occurs at the end of the questionnaire on the mailout form. The overcount question aims to find out if there were any other places where the person could have been duplicated (or potentially should have been counted) on census day.

Between 2004 and 2006, more response categories were added to the question. In both years the question stem read:

Do you sometimes live or stay somewhere else?

In 2004, the response categories were:

To attend college?
To stay at a seasonal or second residence?
To be closer to work?
For a child custody arrangement?
For any other reasons?

In 2006, categories on being in the military, in jail or prison, and in a nursing home were added:

While in the military?
While in jail or prison?
While in a nursing home?

Both census tests showed problematic interviewer behavior and respondent behavior. In 2004, 66 percent of administrations had a major change, and the rate of respondent interruptions was 13 percent (Hunter and Landreth, 2005). The problem cited was that this question has multiple questions embedded within the question. Almost half of the interruptions occurred at first question mark, where respondents thought they could give an answer. Results were similar in 2006 (Childs, Landreth, et al., 2007). In 2006, interviewers were required to read all response options for the first person, and then only needed to read the stem question for later people. However, even for the first person many interviewers only read a subset of the response options.

In cognitive testing, this question was administered in a similar fashion, reading all response options the first time, and just the question stem later times. This did cause some problems for some respondents because they forgot the response options for later people (Childs, Gerber, et al., 2006). Spanish cognitive testing supported this finding, with the researcher commenting that it was "too long and convoluted" (Jones & Childs, 2006).

All of these findings suggest the question was too long. The restructuring of the coverage questions and the residence rules questions allowed us to break this specific question up into several parts. This will, hopefully, improve interviewer administration and respondent comprehension, thus improving coverage.

The other finding from the cognitive testing in both English and Spanish was that the lack of a reference period was problematic (Beck, 2006; Childs, Gerber, et al., 2006). In the revised questions, a reference period is included when we thought it was needed. The following questions were tested in the final round of cognitive testing:

> **a. {Were you/ any of the people you mentioned} living away at college?**
>> Yes – if more than one person in household - **Who?**
>> No
> **b. On April 1st,{were you/ was anyone} living away for the military?**
>> Yes – if more than one person in household - **Who?**
>> No
> **c. On April 1st, {were you/ was anyone} in a place like a nursing home, mental hospital or correctional facility like a jail or prison?**
>> Yes – **Who?**
>> No

Question "a" on college housing correctly identified college students living away from their parental address (Childs, Carter, et al., 2007). It also (correctly) did not identify college students living at home. There was a problem with interviews at a college address that happens to be a regular housing unit (not a dorm or residence hall) and there was some evidence of a problem with no reference period. The recommendation was to add a reference period of "April," rather than "spring semester" so as not to be incongruent with schools that have differing calendars, using quarters or trimesters. Additionally, we recommended probing for people "living in college housing" to eliminate reports of people living in other housing units at college. The wording agreed upon for the 2008 NRFU was:

> **In April, [were you/was anyone] living in college housing?**

Question "b" on military housing was initially tested without a reference period (Childs, Carter, et al., 2007). It became very clear, very quickly that this question was answered drastically differently between different people without a reference period. Since we really are interested in people who were living in military group quarters on census day, it was advantageous to limit the scope of the question to "on April 1st." The question including the reference date correctly identified several people who were living away for the military.

While determining wording for the 2008 NRFU, the reference period was changed from "On April 1st" to "In April" for several reasons. The first was to broaden the group of people flagged for followup due to being away for the military. CFU has additional questions that should define where to count these people in the census. Second, asking a question specifically about April 1

might generate more "don't know" responses. The goal is to cast a fairly wide net for CFU followup. The wording chosen for the 2008 NRFU was:

In April, [were you/was anyone] living away for the military?

Question "c" on other group quarters (nursing home, mental hospital, correctional facility) did not receive any affirmative responses in cognitive testing (we have no way of verifying whether anyone in the interviewed households was in any of these types of places, but we also have no reason to believe respondents were not honest). Generally, this question received a negative reaction. Therefore, we recommended lessoning the emphasis on correctional facilities. Based on the examples used in the residence rules box on the mailout form, the question was shortened to this:

On April 1, 2008, [were you/was anyone] in a place like a nursing home, or a jail or prison?

Additionally, based on cognitive testing, the lead-in phrase of "Now thinking of all the people you just mentioned" was added just prior to the college housing question for households containing more than one person. This will help shift the focus from adding people to the roster to selecting people from the roster.

At the end of the questionnaire, one more overcoverage probe is asked. For the first person, the cognitively-tested question reads:

d. Just to make sure everyone is counted in the right place, during the past year, <did NAME> sometimes live or stay somewhere else to be closer to work, to stay at a seasonal or second residence, to stay with another relative or for any other reason?

For next person:

How about NAME? (Did NAME sometimes live or stay somewhere else for any of those reasons?)
If yes, **For which reason - to be closer to work, to stay at a seasonal or second residence, to stay with another relative or for any other reason?**

Question "d" is asked at the very end of the demographic data collection, as is done in the paper self-administered questionnaire. Initially the plan was to only ask "d" for people who had not already been flagged by one of the previous overcount questions. However, it was decided that everyone should answer "d" to provide consistent data with the mailout form.

In the cognitive test, this question performed fairly well to set a flag for the CFU to more completely determine residence status (Childs, Carter, et al., 2007). Of the fifteen cases where an affirmative response was reported, in 4 cases the person should have been counted elsewhere; in 4 cases it was unclear where someone should be counted (necessitating CFU); and in 7 situations it seemed all people should be counted there (3 of these were movers who stayed elsewhere in a

former home). False positive responses were sometimes due to occasionally staying with a relative, or young adults thinking of their parental home as a "second home." We attribute the false positives of movers reporting yes to the inclusion of a reference period (see Gerber, 2004, for similar problems). Because the mailout form does not have a reference period, and because this test showed that reference period increased false reports, the reference period was dropped from this question for the 2008 NRFU specification. As mentioned before, this question will be asked for each person in the household regardless of whether or not they already have been identified by one of the previous overcoverage questions.

The 2008 question reads as follows:

(Just to make sure everyone is counted in the right place,) did (you/NAME) sometimes live or stay somewhere else such as at a seasonal or second residence, for child custody, or for any other reason?

Addressing mode consistency again, the overcount question from the paper self-administered from was broken up into several shorter questions, which we believe will be more likely to be administered by interviewers. The order of the questions is different than on the mailout, but this is done to eliminate the repetitiveness of the residence rules and the overcount question. We believe this will improve the consistent delivery of all the information.

The Housing Question

Tenure Question

Consistently, the tenure question has been one of our most challenging to adapt to the interviewer-administered mode. In 2004 and 2006, the question read identically:

Is this (house / apartment / mobile home)…
> **Owned by you or someone in this household with a mortgage or loan?**
> **Owned by you or someone in this household free and clear?**
> **Rented for cash rent?**
> **Occupied without payment of cash rent?**

In both tests, it suffered from the highest rate of major change of any question in the NRFU (67 percent and 64 percent for 2004 and 2006, respectively; Hunter & Landreth, 2005; Childs, Landreth, et al., 2007). In most cases, interviewers omitted one or more of the response options or simply asked if the place was "owned or rented." Respondents often had difficulty answering this question as well. Sometimes respondents answered in similar (though not the same) terms, saying things like "I'm buying it" or "It's deducted from my check." Other times the respondent misinterpreted the intent of the question and responded with the name of the person who owns it. At the end of the interaction, it was still sometimes unclear for owners whether the person owned with a mortgage or owned free and clear.

In cognitive testing, the question often needed to be repeated for understanding in English and Spanish (Hunter, 2005; Jones & Childs, 2006). The term "cash rent" caused problems in both

English and Spanish as well (Childs, Gerber, et al., 2006; Hunter, 2005; Jones & Childs, 2006; Omsted & Hourcade, 2005).

For this question, we recommended using a technique called "unfolding," that is, administering shorter questions with tailored follow-up probes that can be administered easily on an automated instrument, but that would be too difficult for implementation on paper due to the complex skip patterns (see Fowler et al., 1999). During the final round of cognitive testing, we tested an unfolding question, which read like this:

Is this house owned by you or someone in this household?
    Yes – **Is it owned with a mortgage or owned free and clear?**
    No – **Is it rented?**

Respondents also had difficulty with this question, focusing on the "who" aspect of the question (e.g., do you own it or does someone else?; Childs, Carter, et al., 2007). Based on this finding and on recommendations of the Mode Consistency Guidelines (Martin et al., 2007), the question in the 2008 NRFU specification mimics the same question used in the American Community Survey and reads like this:

Do you or does someone in this household own this <house/apartment/mobile home> with a mortgage or loan (including home equity loans), own it free and clear, rent it or occupy it without having to pay rent?

"House," "apartment," or "mobile home" will now be an automated fill that the interviewer chooses in the beginning of the interview. This will prevent one of the most common problems faced by the mailout version of this question, which is respondents thinking that the question is asking if the place is a house, apartment, or mobile home.

In future testing, we recommend the continued testing of an unfolding question, this time focusing on owning and renting. Here is a suggestion:

Do you or does someone in this household own this <house/apartment/mobile home> or do you rent it?
    Own - **Is it owned with a mortgage or owned free and clear?**
    Rent
    Neither - Occupied without payment of rent

We also recommend removing the phrase "including home equity loans" because it is unclear how and when to administer that phrase (parentheses are used conventionally for "read if necessary" text). Instead we suggest adding this sentence as optional text (to be read if the respondent has questions or demonstrates confusion): *For the purposes of this Census, home equity loans are considered mortgages.*

Determining Relationships to the Householder

Reference Person

Recall that in the section on gathering a roster, we recommended separating the tasks of gathering a roster from determining the reference person (or householder). In the final round of cognitive testing, we tested a separate question to determine the reference person, which followed the tenure question (Childs, Carter, et al., 2007). The Census Bureau defines the reference person (or householder) as the person who owns or rents the housing unit. The tested question read as follows:

> **What is the name of the person or one of the persons who (owns/rents) this (house/apartment/mobile home)?**

This question fills the appropriate term (owns or rents) based on the answer to the tenure question.

Cognitive testing respondents had problems with this question, answering with the mortgage company or the landlord, rather than a resident owner (Childs, Carter, et al., 2007).

We also tested formalizing the typically informal rules for what to do with multiple owners or renters. If the respondent was identified as one of the owners or renters, we used that person as the reference person. If not, we used the first person that the respondent identified. For testing purposes, we also asked the respondent to choose a single householder. Our implemented rules almost always led to a female reference person, whereas the forced choice led to a male reference person (Childs, Carter, et al., 2007). Because of this difference, we did not recommend implementing these rules in an automated system.

The recommendation from this testing and recommendations from the Mode Consistency Guidelines (Martin et al., 2007) led to the wording that will be used in the 2008 NRFU:

> If the housing unit is owned:
> **Of the people who live here, who owns this <house, apartment, mobile home>?**
>
> If the housing unit is rented:
> **Of the people who live here, who rents this <house, apartment, mobile home>?**

If the housing unit is occupied without payment of rent, then the resident respondent is the reference person.

We also recommended adding an optional instruction:

> **I can only record one name.**

This was not implemented in the 2008 specification, but something like this was also recommended by the contractor's usability testing.

Relationship Question

During the testing cycle, the relationship question has caused problems on several fronts. First, the HHC has a small screen, which led to difficulty fitting all the response categories on the screen. In order to deal with this, in 2004, the programmers implemented a "blind branching," where the interviewer presented respondents with a flashcard and asked this question:

Which one of these categories best describes how [you are/NAME is] related to [you/NAME]?

The flashcard displayed all the response categories (related and not related categories) to the respondent, but the interviewer had to input on the first screen either "yes, related" or "no, not related" to get the appropriate list of categories to select from. This caused the interviewer to often change the question to the respondent to first ask whether the two people were related or not (Hunter & Landreth, 2005).

In 2006, branching scripted and respondents were first asked if the two people were related:

> **[Are you/is NAME] related to [NAME]?**
> > Yes – Go to a
> > No – Go to b

And then, based on the answer, the specific relationship was identified:

> **a. Which one of these categories best describes how [you are/NAME is] related to [NAME]?**
> > Husband or wife
> > Biological son or daughter
> > Adopted son or daughter
> > Stepson or stepdaughter
> > Brother or sister
> > Father or mother
> > Grandchild
> > Parent-in-law
> > Son-in-law or daughter-in-law
> > Other relative

> **b. Which one of these categories best describes [your/NAME's] relationship to [NAME]?**
> > Roomer, boarder
> > Housemate, roommate
> > Unmarried partner
> > Foster child or foster adult
> > Other nonrelative

The branched related-or-not-related question was shown to be very problematic through cognitive testing. Respondents often do not categorize relationships like the Census Bureau does. Many respondents report that spouses are not related to each other (Hunter, 2005). Other

respondents in both English and Spanish cognitive testing disagree with how we categorize other relations, including foster children, adopted children, and unmarried partners (Beck, 2006; Hunter, 2005; Jones & Childs, 2006). If a respondent answers the first question "incorrectly" by saying his spouse is not related to him, then the interviewer will get the wrong set of response options to the followup, possibly inducing the interviewer to select the wrong option instead of going backwards.

Cognitive testing also demonstrated reporting reverse relationships (Childs, Gerber, et al., 2006), a problem also noted during the 2004 Census Test production data (Love & Byrne, 2005). For example, if the householder is 35 years old, and the householder's child, Mary, is 10 years old, the question may appear "How is Mary related to you?" The respondent may answer by saying "I'm her mother" or "that's my daughter." In this situation, the answer the Census Bureau expects to be recorded is "son or daughter" because the relationship of Mary to the householder is child. If the interviewer records "mother or father" it will cause an error in the data which appears as though the 10-year-old is the mother of the 35-year-old. This type of error occurred in dramatic rates during the 2004 Census Test – as many as 60 percent of parents reported were more than 10 years younger than the reference person, indicating a reversal in the relationship reporting (Love & Byrne, 2005).

Other respondents in cognitive testing incorrectly reported all relationships to him or herself, despite the fact that the reference person was someone else (Childs, Gerber, et al., 2006). It is a cognitively difficult task to report relationships to someone other than the person's self (for an anthropological discussion of ego-centered kin terms, see Fox, 1991). We attribute both of these problems to the complex question that does not necessarily imply direction. The other finding from behavior coding was that often respondents only report "son or daughter" and do not specify whether the child is biological, adopted or step (Childs, Landreth, et al., 2007).

Based on these findings, the last round of cognitive testing tested a novel idea of a "fill-in-the-blank" format for the relationship question. This would be more directionally indicative and hopefully solve the problem of reversals of relationship reporting. Additionally, we entirely omitted the screener "related" question. This is the question we tested:

> **Next I need to record relationships of everyone to (REFERENCE PERSON). Using the categories on the card, please help me fill in the blanks.**
>
> **NAME is (REFERENCE PERSON)'s _____.**
>
> Husband or wife   Roomer or boarder
> Biological son or daughter   Housemate or roommate
> Adopted son or daughter   Unmarried partner
> Stepson or stepdaughter   Other nonrelative
> Brother or sister
> Father or mother
> Grandchild
> Parent-in-law
> Son-in-law or daughter-in-law

Other relative

This method was very successful in the cognitive testing (Childs, Carter, et al., 2007). We did not see any evidence of inverted relationships. The subject matter area was very supportive of this finding, and we hope to test it in a field test in preparation for the 2020 Census. There could be issues with translating this question into Spanish – research is yet to show how to best do that.

Because this experimental question is still being studied, the 2008 NRFU question is scripted to not use a screener, but to ask the original relationship question (worded more closely to the mailout form than it had been in 2004 and 2006, hopefully to improve directional reporting). In a personal visit, the relationship question will appear like this:

> **Please look at List A on the handout I gave you at the beginning of the interview. How is [NAME] related to [PERSON 1'S NAME]?**
>
> Husband or wife   Roomer or boarder
> Biological son or daughter   Housemate or roommate
> Adopted son or daughter   Unmarried partner
> Stepson or stepdaughter   Other nonrelative
> Brother or sister
> Father or mother
> Grandchild
> Parent-in-law
> Son-in-law or daughter-in-law
> Other relative

The flashcard can be shown in the personal visit, to increase reporting of type of child. Because there is no flashcard in the telephone mode, the telephone version is somewhat different. The interviewer is not required to read all response categories to the respondent (due to the long list). The question is, instead, field-coded by the interviewer. If a response of "son or daughter" is given, then a followup question determines whether the child is biological, step or adopted. The script reads as follows:

> **How is [NAME] related to [PERSON 1'S NAME]?**
> Husband or wife
> Son or daughter → **Is [NAME] [PERSON 1's] biological son or daughter, adopted son or daughter, OR stepson or stepdaughter?**
>   Biological son or daughter
>   Adopted son or daughter
>   Stepson or stepdaughter
>   Foster son or daughter
> Brother or sister
> Father or mother
> Grandchild
> Parent-in-law
> Son-in-law or daughter-in-law
> Other relative

> Roomer or boarder
> Housemate or roommate
> Unmarried partner
> Other nonrelative

An additional feature of the followup probe is that it allows reporting of foster children, but codes them back to "other nonrelative" since there is no longer a response category for foster children. We automated this option and back coding because we think respondents will report foster children as sons or daughters, and interviewers will need to be able to record something on the followup screen. Notice that the foster child category is not read aloud to respondents, but is blind-coded.

Another addition to the relationship question series for 2008 is to implement a soft edit after the age screen to deal with seemingly inverted relationships. After the age has been determined for all household members, a relationship edit check is conducted to look for household members who are reported as a parent or parent-in-law to the reference person, yet their age is younger than the reference person. The edit reads:

I have recorded that NAME is NAME's [parent/parent-in-law]. Is that correct?

This question will catch the most predominant errors, which are inversions in that direction. However for 2010, I also recommend adding the inverse, when a biological or adopted child is reported that is older than the reference person, we should verify:

I have recorded that NAME is NAME's [child]. Is that correct?

These questions were field tested in a Random-Digit-Dial research study conducted by Nichols and Childs (2007). Results were rather inconclusive because it was not triggered in many instances. The authors attribute that to the fact that the interviewers were permanent Census Bureau telephone interviewers who are used to administering our relationship question, as opposed to novice interviewers like those used for decennial operations.

Sex

The sex question is one of the few questions for which the script has not changed in the NRFU instrument between 2004 and 2008. However, the interviewer instructions have changed. In 2004 and 2006, interviewers were not supposed to verify sex. However, behavior coding noted that the question was skipped 48 percent of the time in 2004 (Hunter & Landreth, 2005). We attribute this to the awkwardness of asking for sex when the context of the interview has clearly indicated that the person in question is of a specific sex. Cognitive testing also noted the awkwardness of asking this question to respondents when the answer is apparent. Based on these findings, the 2008 NRFU specification allows the interviewer to verify sex. The specific, on-screen instruction is rather complex:

> Ask or verify sex for all persons. Do not enter/record sex based upon observation or responses to other items.

ARE YOU/IS NAME male or female?

In my opinion the interviewer instruction is too long for an instruction that appears on the small handheld screen. That type of advice is better suited for training materials. The screen should have a simple reminder like "Ask or verify." This is standard survey practice.

Age and Date of Birth

In 2004, the age question appeared before the date of birth question in the NRFU, mimicking the mailout form. This structure failed to take advantage of the automated capability to verify age given date of birth, rather than asking for both. Spencer and Perkins (1998) recommended accepting age calculated from birthdate when possible. When using an automated instrument, the interviewer can verify calculated age rather than asking both questions.

In 2006, the order of the questions was switched. Date of birth was asked first, and age was calculated as of April 1st and verified as of that date (in the past). This reordering is in line with recommendations from the Mode Consistency Guidelines (Martin et al., 2007).

> **a. What is YOUR/NAME'S date of birth?**
> > b. DK- **What was YOUR/NAME'S age on April 1, 2006? (If you don't know the exact age, please estimate.)**
>
> **c. For the census, we need to record age as of April 1, 2006. So, just to confirm – NAME was AGE on April 1, 206?**

The predominant problems with this question noted through behavior coding and cognitive testing were unnecessarily repeating the sentence "For the census, we need to record age as of April 1, 2006" and a programming error that caused estimated ages to be confirmed (Childs, Landreth, et al., 2007; Childs, Gerber, et al., 2006). Two recommendations were made to address these issues: 1) only read the introduction sentence for the first person for whom age is verified; and 2) only confirm age when date of birth is given. These were both accepted for the 2008 NRFU, and the questions will be as follows:

> **a. What is YOUR/NAME'S date of birth?**
> > b. DK- **What was YOUR/NAME'S age on April 1, 2008?**
> > (If you don't know the exact age, please estimate.)
> >
> > Make sure the respondent gives the age in completed years as of April 1, 2008. Do not round up.
> > Do not enter age in months.
> > For babies less than 1 year old enter 0 as the age.
>
> **c. (For the census, we need to record age as of April 1, 2008.) So, just to confirm – NAME was AGE on April 1, 2008?**

I have a few notes on areas I would like to see improved as we move towards 2020. As with the sex question, I think there are excessive interviewer instructions on the screen. Most of these are issues that should be covered in training, and not left to clutter the small handheld screen.

The larger issue is that there were indications from behavior coding and cognitive testing that reporting or verifying a date in the past was burdensome for interviewers and respondents. The Census Bureau needs to record age as of census day, which is April 1st. When asking for age as of April 1st, interviewers sometimes left off the date – changing the question to a current age question (Hunter & Landreth, 2005). Additionally, interviewers sometimes changed the age verification question to a present-tense question, confirming current age instead of census day age (Childs, Landreth, et al., 2006). Both of these could lead to incorrect reports.

In the 2000 Census, Carter and Brady (2002) found two situations where respondents had problems reporting age correctly on the Census form – both the self-administered form and the interviewer-administered paper NRFU form, which asked for age as of April 1st. The most notable problem identified was that respondents misreported age when the person's birthday was after April 1st and also after the date that the form was checked-in (meaning the person likely completed the form after his or her birthday, which made his or her current age older than his or her age on April 1st). Forty percent of the people in this category over-reported their age (suggesting they reported their current age, and not their age as of April 1st.) The other problem occurred when the person's birthday was before April 1st, and their form was checked-in before April 1st. In this situation, the person should report their age as of a date in the future. In 10.3 percent of cases, the person underreported his or her age, indicating that they were reporting their current, not future, age on April 1st.

Nichols and Childs (2007) make the argument that, when possible, current age should be verified. Given a complete date of birth, age as of census day can be stored, while calculating and verifying current age with the respondent – a much easier task. In a small experiment conducted by Nichols and Childs, there was some indication that interviews that verify age in the past will take longer, cause more interviewer and respondent burden, and consequently cost more, than interviews that verify current age.

We recommend testing, leading up to the 2020 Census, confirming current age when date of birth is given. If date of birth is not given, age could still be asked as of census day.

Origin and Race

Hispanic Origin

In 2004, we tested a two-part Hispanic origin and Race series. Both questions had a series of followups to gather detailed origin and race data. In 2006, we tested a three-part series, with a yes/no Hispanic origin question, a race question that elicited selection of one of the five major race categories (or some other race) and an ancestry question that gathered detailed origin data for everyone.

The only difference between the initial Hispanic origin questions in 2004 and 2006 was a change in the order in which the terms in the question were presented. This was based on testing done with the mailout form, not the NRFU (Gerber & Crowley, 2005). In 2004, the series was: Spanish, Hispanic or Latino origin. In 2006 and later, the order is: Hispanic, Latino or Spanish origin.

The 2006 and 2008 versions of the question read like this:

[ARE YOU / IS NAME] of Hispanic, Latino or Spanish origin?

Many people interpret the Hispanic origin question to be a multiple choice question rather than a yes/no question (Beck, 2006; Childs, Landreth, et al., 2006; Jones & Childs, 2006). This causes some unnecessary respondent burden, but does not necessarily impact data quality. Hispanic respondents often provide a nationality to the Hispanic origin question (instead of answering "yes"). This is only problematic if the interviewer does not know whether the origin mentioned is a Hispanic origin. We witnessed an example of this during the 2006 behavior coding where the interviewer verified with the respondent that "Mexican" was not "Hispanic, Latino or Spanish origin" (Childs, Landreth, et al., 2007). Though this is a dramatic example, there are many less known Hispanic countries that most interviewers would likely not recognize or be able to place (e.g., Brazil, which is a Latino country, but not a Spanish-speaking country, and thus non-Hispanic). Non-Hispanic respondents in the cognitive test sometimes asked whether certain nationalities were considered Hispanic (e.g., Cuban and Italian; Hunter, 2005). This could be problematic for reporting for other household members. Finally, a few respondents interpreted this question as citizenship question which could cause privacy concerns (Childs, Carter, et al., 2007).

In 2004, a Hispanic origin followup question was asked of anyone who answered "yes" to the initial Hispanic origin question. This was an adaptation made to the mailout version of the questionnaire, which embedded answers to a second question into the response set of the original Hispanic origin question.

2004 Mailout Hispanic origin Question
Is Person 1 of Spanish, Hispanic or Latino origin?
[] No, not of Spanish, Hispanic or Latino origin
[] Yes, Mexican, Mexican American, or Chicano
[] Yes, Puerto Rican
[] Yes, Cuban
[] Yes, Another Spanish, Hispanic, or Latino origin, for example, Argentinean, Columbian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on. *Print origin.*

2004 NRFU Hispanic origin Followup Question
(Are you/Is NAME) Mexican, Mexican American, or Chicano? Puerto Rican? Cuban? Another Spanish, Hispanic, or Latino origin? (For example, Argentinean, Columbian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)

This question was followed by a second, open-text question in the event that an "other" option was chosen from this first screen. The 2004 Hispanic origin format is similar to the 2008 NRFU format.

Hispanics, and particularly Spanish-speaking Hispanics, often do not identify as "Hispanic, Latino or Spanish origin" (Childs, Landreth, et al., 2007). This is demonstrated by the high rates of reporting nationalities to the Hispanic origin question in the 2006 behavior coding (39% of Spanish-speaking respondents reported with a nationality rather than answering affirmatively). The risk is, as mentioned above, if interviewers do not recognize the nationality as being Hispanic, then the interviewer might input the incorrect answer. Additionally, we do not know how many respondents may answer "no" to this question incorrectly because they do not know their country of origin is among those considered "Hispanic." Because the initial question is a yes/no only, there is some risk that interviewers and respondents will not understand what is meant by "Hispanic, Latino or Spanish origin." For this reason, it was decided to use a flashcard for this question that presents the response categories as they appear in the mailout form. This provides the respondents (and interviewers) with the same information provided to respondents in the self-response mode, as recommended by the Mode Consistency Guidelines (Martin et al., 2007). Thus, in 2008 when answering the initial Hispanic origin question, respondents will see a list on their information sheet that looks like this:

Hispanic, Latino, or Spanish Origin

> **No, not of Hispanic, Latino, or Spanish origin**
> **Yes, Mexican, Mexican American or Chicano**
> **Yes, Puerto Rican**
> **Yes, Cuban**
> **Yes, Another Hispanic, Latino or Spanish origin,** *For example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard and so on*

The followup question text will read all the answer categories as well as the examples. The text of the two questions will be as follows:

> **Please look at List B. Are you of Hispanic, Latino or Spanish origin?**

> If Yes, ask **Are you Mexican, Mexican American, or Chicano; Puerto Rican; Cuban; or of another Hispanic, Latino or Spanish origin; for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on?**

Race

In the 2004 Census Test, the full race question (with 14 race categories) was tested, with followup questions for detailed American Indian or Alaska Native tribe, other Asian or other Pacific Islander races. In 2006, a shortened 5-category race question was tested, paired with the ancestry question.

The questions were as follows:

2004
> **Using this list, please choose one or more races that you consider yourself to be. White, Black, African American or Negro, American Indian or Alaska Native, Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, Other Asian (for example: Hmong, Laotian, Thai, Pakastani, Cambodian), Native Hawaiian, Guamanian or Chamorro, Samoan, Other Pacific Islander (For example: Fijian, Tongan)**

2006
> **What is [YOUR/NAME'S] race? (You may choose one or more races. For this census, Hispanic origins are not races. White or Caucasian, Black, African American or Negro, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Some other race?)**

In the 2004 behavior coding, one of the major problems identified was that interviewers skipped the question (Hunter & Landreth, 2005). After the first time it was administered for the first person in the household, it was skipped for almost half of all administrations. This indicates that either respondents told the interviewers that all household members were of the same race (and the interviewer did not verify for each person) or that interviewers assumed all household members were of the same race. These are vastly different problems. The first is merely a failure by the interviewer to verify previously reported information. The latter is equivalent to falsifying data (i.e., entering data that was not reported by the respondent).

In 2006, when the interview was topic-based, rates of skipping the race question were slightly lower (Childs, Landreth, et al., 2007). After it was administered the first time, it was skipped 30 percent of the time for later household members. This still indicates a fundamental problem with interviewer behavior that should be remedied with better training. Interviewers also made a variety of changes to the question wording. For the first person, the entire question (including the parenthetical statement) was required reading. For later household members, only the initial question was required reading. However, interviewers only read the question in full for the first person 34 percent of the time. In other cases, they omitted all, or some, of the response options, or omitted the phrase "Hispanic origins are not races."

In both behavior coding studies, respondents frequently gave inadequate answers. In 2004, only 42 percent of respondents initially gave adequate answers, and in 2006, only 38 percent did (Hunter & Landreth, 2005; Childs, Landreth, et al., 2007). In both cases, most inadequate responses involved giving some variant of Hispanic origin as a race – either by giving a nationality, one of the terms in the Hispanic origin question, or a term like "trigueña" or "mestiza."[11] This was consistent with cognitive test findings (Beck, 2006; Childs, Gerber, et al., 2006; Hunter, 2005; Jones & Childs, 2006). Additionally, respondents in cognitive testing did not attend to the instruction to choose one or more races, possibly because it is buried with other information (Childs, Gerber, et al., 2006; Hunter, 2005; Jones & Childs, 2006). The statement "Hispanic origins are not races" was neither understood nor followed by respondents (Childs,

---

[11] Trigueña is a Spanish term used to refer to skin color, meaning "olive skinned." Mestiza.is a Spanish term for mixed race, particularly of European and indigenous origin.

Gerber, et al., 2006). In fact, a couple of respondents even misinterpreted the phrase to say "Hispanic origins are not racists" – a misinterpretation that we think is a critical defect in the question wording. Additionally, some respondents in cognitive testing found the word "negro" to be outdated and/or offensive (Childs, Carter, et al., 2007).

Because of the OMB guidance regarding data collection of race and Hispanic origin data, the basic race categories stand firm. Because of concerns about mode consistency, the Hispanic origin note also remains unchanged. However, in the 2008 NRFU specification, some improvements were made.

For Person 1, the following text will be read:

Please look at List C on the handout I gave you at the beginning of the interview. You may choose one or more races. (For this census, Hispanic origins are not races.) Are you White; Black, African American, or Negro; American Indian or Alaska Native; Asian; Native Hawaiian or Other Pacific Islander; or Some other race?

The parenthetical note about Hispanic origins will be read in households where at least one member is reported to be Hispanic.

For Person 2, the following will be read:

Is Person 2 White; Black, African American, or Negro; American Indian or Alaska Native; Asian; Native Hawaiian or Other Pacific Islander; or Some other race?

For Persons 3 and higher, the question will be shortened[12]:

What is Person 3's race? (Is Person 3 White; Black, African American, or Negro; American Indian or Alaska Native; Asian; Native Hawaiian or Other Pacific Islander; or Some other race?)

Immediately following a report of "American Indian or Alaska Native," "Asian," "Native Hawaiian or Other Pacific Islander," or "Some other race," a followup question will be asked. The parenthetical text below indicates wording that will appear as required reading the first time the question is presented in a household. The followup questions are as follows:

American Indian: **(You may list one or more tribes.)What is your enrolled or principal tribe?**

Asian: **(You may choose one or more Asian groups.) Are you Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese or some other Asian group, (for example Hmong, Laotian, Thai, Pakistani, Cambodian and so on)?**

---

[12] The exact text depends on the response given for the previous household member. If no followups are required to the previous household member's initial race response, the question wording is shortened to "How about NAME?" with the remainder in parenthesis as shown above.

Other Asian group: **What is that other Asian group?** (For example Hmong, Laotian, Thai, Pakistani, Cambodian and so on)

Native Hawaiian or Other Pacific Islander: **(You may choose one or more Pacific Islander groups.) Are you Native Hawaiian; Guamanian or Chamorro; Samoan; or another Pacific Island group, (for example Fijian, Tongan and so on)?**

Other Pacific Island group: **What is that other Pacific Islander group?** (For example Fijian, Tongan and so on)

Some Other Race: **What is your other race group?**

Topic-based administration of Hispanic origin and Race

There was some discussion of whether to gather detailed Hispanic origin and race data immediately after the screener question for an individual, or whether to gather the initial response for the whole household, and then later gather the detailed information for each relevant person (e.g., asking whether each household member is Hispanic, then gathering detailed origin for each person reported as Hispanic). Cognitive testing that happened concurrently with the NRFU and American Community Survey demographic questions tested each approach separately. In the final round of NRFU cognitive testing (Childs, Carter, et al., 2007), for the Hispanic Origin question, the followup country of origin question was asked immediately after receiving a "yes" response to the initial question. This flowed something like this:

I: Are you of Hispanic, Latino or Spanish origin?
R: Yes
I: *Show flashcard.* Are you Mexican, Mexican American, Chicano, Puerto Rican, Cuban, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, or of another Hispanic, Latino or Spanish origin?
R: I'm Cuban.
I: How about Mary? Is she of Hispanic, Latino or Spanish origin?
R: No
I: How about Juana? Is she of Hispanic, Latino or Spanish origin?
R: Yes, she's Dominican.

The only difficulty noted with this type of administration was with the implementation of the flashcard for the followup question. Each time a "yes" response was recorded to the initial question, the followup was immediately asked. Since the followup had a flashcard, this presented some technical difficulties for the interviewer needing to direct attention to and away from the flashcard between asking about different household members.

The Race question was administered similarly in this cognitive test (Childs, Carter, et al., 2007). A flashcard showing all race response options was shown to the respondent initially, and then detailed race was captured immediately. Here is an example of how this might have sounded:

I: *Show flashcard.* Using this list, please choose one or more races that you consider yourself to be. White or Caucasian, Black, African American, or Negro, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, or Some other race?
R: I'm American Indian.
I: What is your enrolled or principal tribe?
R: Cherokee
I: *Show flashcard* And how about Jack?
R: Jack is White.
I: And how about Lydia?
R: She's Asian and Pacific Islander.
I: Which of the following Asian groups is she from? Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese or some other Asian group such as Hmong, Laotian, Thai, Pakistani, Cambodian and so on?
R: She's Thai.
I: Which of the following Pacific Islander groups is she from? Native Hawaiian, Samoan, Guamanian or Chamorro, or some other Pacific Island group such as Fijian, Tongan and so on?
R: I think she's Fijian.

This method of followup worked well during cognitive testing.

Chan (2006) tested an alternate way of administering followup questions for both Hispanic origin and race. Chan asked the initial question for all household members first, using a typical "how about" structure after the question was asked in full for the first person. Then, Chan asked the relevant followup questions for each household member for which it applied. This would sound something like this:

I: Are you of Hispanic, Latino or Spanish origin?
R: Yes
I: How about Mary?
R: No
I: How about Juana?
R: Yes.

I: *Show flashcard.*
Are you Mexican, Mexican American, Chicano, Puerto Rican, Cuban, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, or of another Hispanic, Latino or Spanish origin?
R: I am Puerto Rican.
I: How about Juana?
R: She is Dominican.

I: *Show flashcard* What is your race? Are you White or Caucasian, Black, African American, or Negro, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, or Some other race?
R: I'm American Indian.

I: And how about Jack?
R: Jack is White.
I: And how about Lydia?
R: She's Asian and Pacific Islander.

I: What is your enrolled or principal tribe?
R: Cherokee

I: Which of the following Asian groups is Lydia from? Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese or some other Asian group such as Hmong, Laotian, Thai, Pakistani, Cambodian and so on?
R: She's Thai.

I: Which of the following Pacific Islander groups is Lydia from? Native Hawaiian, Samoan, Guamanian or Chamorro, or some other Pacific Island group such as Fijian, Tongan and so on?
R: I think she's Fijian.

Based on joint meetings between survey methodologists and subject matter experts, it was decided that for telephone and personal-visit interviews, Hispanic origin and race questions would be branched using an immediate follow-up approach, as was tested in the NRFU cognitive testing.

Additionally, as noted before, it was decided that for the Hispanic Origin question the flashcard should be presented with the first Hispanic origin (yes/no) question so that respondents would understand what nationalities are considered Hispanic. The initial Hispanic origin question will be asked while the respondent sees the flashcard in personal-visit interviews, and then immediately afterwards detailed origin will be gathered. This will facilitate reporting because respondents often immediately provide their country of origin to the Hispanic origin question (see previous discussion).

For the race question in the 2008 NRFU script, interviewers will first read the six main race categories to the respondent. If the response(s) to the main question require follow-up (American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, or Some other Race), the interviewer will administer the follow-up(s) for that person immediately. Example statements will be read in the appropriate follow-ups. In personal-visit interviews, a flashcard will be used that lists all of the detailed race categories and examples. Thus, the respondent may provide either a high-level or detailed response at the time the initial race question is presented. By structuring the interview such that we perform an immediate follow-up for the appropriate categories, the interviewer can easily record either a high-level or detailed response. This question structure and use of the flashcard worked effectively during the Childs, Carter, et al. (2007) cognitive testing.

Ancestry

Because the ancestry question is no longer being used, I will not review the findings from the 2006 Census Test on this question. Please see the individual reports if you are interested (Beck, 2006; Childs, Gerber, et al., 2006; Childs, Landreth, et al., 2007; Jones & Childs, 2006).

Conclusions

We saw evidence of non-standardized interviewing from behavior coding using the 2004 and 2006 instruments. We saw evidence of respondent confusion and burden with long questions from cognitive testing using the 2006 instrument. However, we believe that a more standardized interview can be achieved if we fully exploit automation. This was the goal behind the development of the revised 2008 NRFU. See Appendix B for a somewhat simplified example of the 2008 NRFU script. We believe that this interview will lead to a higher data quality than the previous versions.

Because we had evidence that the interviewers did not do well with the roster and coverage questions, we developed a series of shorter questions to convey residence rules. Based on survey research and evidence from the studies presented here, we were able to use a topic-based structure to shorten questions where it was appropriate. We created an automatic fill for "house," "apartment," or "mobile home" based on interviewer observation. All of these changes should result in a 2008 questionnaire that is less burdensome for respondents, as well as interviewers. Additionally, by adhering to the Mode Consistency Guidelines, the data should be consistent with that which would have been gathered by the paper form. Data quality should be improved from past census tests, due to the edit implemented in the relationship question, as well as the improvements to the residence rules presentation.

These revisions will be evaluated during the 2008 Dress Rehearsal through a study that will combine observations, expert respondent debriefings and behavior coding of the English and Spanish questionnaires. For a case study that demonstrates how this evaluation will be carried out, see Nichols and Childs (forthcoming). Staff with survey methodology and subject matter expertise will go into the field during the NRFU dress rehearsal to observe interviews, tape record those interviews for subsequent behavior coding, and conduct respondent debriefings in situations where the expert identifies problems with the NRFU interview. This method will let us fully and quickly understand how the revisions made to the NRFU are working in a field environment and allow us to make any necessary changes before the 2010 Census.

Future Research

Originally, during the 2004 test, questionnaire wording had been optimized for a self-administered paper form and applied to an interviewer-administered CAPI instrument. Research described here demonstrated that when question wording is designed with only a single mode in mind, using the exact same wording across modes can lead to detrimental effects on interviewer behavior. During development for the 2008 Dress Rehearsal, the Census Bureau used the "customization by mode" approach to multimode questionnaire development. This model entails having question wording optimized for each mode (and, consequently, question wording that

may differ by mode) rather than having wording that is a compromise between wording that would work best in each mode. In contrast, Dillman and Christian (2005) advocate a "unimode" approach to survey design with multimode instruments that involves, instead, developing a single question wording considering all survey modes that will be used. Both the unimode and the customization by mode models are intriguing and should be empirically tested to examine effects on response distributions and ease of administration (see the recent discussion of this topic in de Leeuw 2005).

Acknowledgements

References Cited

Beck, J. (2006). Cognitive test of the 2006 Spanish NRFU: Round 1. Internal Report.

Bosley, J., Conrad, F. G., and Uglow, G.. (1998). Pen CASIC: Design and usability. In M. P. Couper, R. P. Baker, J. Bethlehem, C.Z.F. Clark, J. Martin, W. L. Nicholls, II, and J. M. O'Reilly, eds., Computer-assisted survey information collection, 521–41. New York: John Wiley.

Cameron, M. (2005). Bleeding edge or proven technology? The fact and fiction of mobile survey computing. Paper presented at the Association for Survey Computing's Conference on Mobile Computing, April 22.

Cannell, C., Fowler, F., and Marquis, K. (1968). The influence of interviewer and respondent psychological and behavioral variables on reporting in household interviews. Vital and Health Statistics, Series 2 (26). Washington, DC: U.S. Government Printing Office.

Chan, A. (2006). Report on Topic-Based Mode Consistency CATI/CAPI Cognitive Prestesting in the American Community Survey (ACS). Internal draft report, August 14, 2006.

Childs, J. H., Carter, III., G. R., Norris, D., Hanaoka, K., and Schwede, L. (2007). Cognitive Test of the NRFU Round 3: Revised Questions. SRD Research Report Series RSM2007/09.

Childs, J. H., Gerber, E. R., Carter, G., and Beck, J. (2006). Cognitive Test of the 2006 NRFU: Round 2. SRD Study Series Report SSM2006/05.

Childs, J.H. and Landreth, A. (2006). Analyzing Interviewer/Respondent Interactions while Using a Mobile Computer-Assisted Personal Interview Device. Field Methods, Vol. 18, No. 3, 335-351.

Childs, J. H., Landreth, A., Goerman, P., Norris, D. and Dajani, A. (2007). BEHAVIOR CODING ANALYSIS REPORT: Evaluating the English and the Spanish Versions of the Non-Response Follow-Up (NRFU) for the 2006 Census Test. SRD Study Series Report SSM2007-16.

Colosi, R. (2001). An Analysis of Time Stamp Data: Person-based Interviewing vs. Topic-based Interviewing. Unpublished report.

Couper, M. P. (2002). New technologies and survey data collection: Challenges and opportunities. Paper presented at the International Conference on Improving Surveys; Copenhagen, 25-28 August.

de Heer, W. F. (1991). The use of handheld computers in social surveys of the Netherlands Central Bureau of Statistics. The Statistician 40:125–38.

de Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. Journal of Official Statistics 21:233–55.

Dillman, D. A., and Christian, L. M.. (2005). Survey mode as a source of instability in responses across surveys. Field Methods 17 1: 30–52.

Dumas, J. S. and Redish, J. C. (1999). A Practical Guide to Usability Testing. 1st. Intellect Books

Fowler, F. (1992). How unclear terms affect survey data. Public Opinion Quarterly 56:218–31.

Fowler, F. J., Fitzgerald, G., and Beatty, P. (1999). When two questions are better than one. 1999 Proceedings of the Section on Survey Research Methods, American Statistical Association, Alexandria, VA: American Statistical Association, 956-61.

Fox, R. (1991). Encounter with Anthropology. Edison, NJ: Transaction Publishers.

Gerber, E. R. (1994). The Language of Residence: Respondent Understandings and Census Rules. Working Paper in Survey Methodology no. 94/05, Census Bureau.

Gerber, E., (2004). Summary Report of Cognitive Testing on Residence Rules Instructions. Internal Report.

Gerber, E., and Crowley, M. (2005). Report on Cognitive Testing of a Shortened Sequence of Hispanic Origin, Race, and Modified Ancestry Questions: Content development for the 2005 National Content Test. Internal report. October 18, 2005.

Goerman, P. (2005). Making ends meet: The complex household as a temporary survival strategy among new Latino immigrants to Virginia. In Shwede, L., Blumberg, R. L., and Chan, A. Y. (Eds). Complex Ethnic Households in America. Lanham, MD: Rowman & Littlefield Publishers, Inc. 149-180.

Gravlee, C. C. (2002). Mobile computer-assisted personal interviewing (MCAPI) with handheld computers: The Entryware System v3.0. Field Methods 14 3: 322–36.

Greene, P. D. (2001). Handheld computers as tools for writing and managing field data. Field Methods 13 2: 181–97.

Heimel, S., King, R., and Sheppard, D. (2007). 2006 Census Test Evaluation #1: The Residence Rule and Residence Instructions. DSSD 2006 Census Test Memorandum Series #I-14, September 11, 2007.

Hunter, J. (2005). Cognitive Test of the 2006 NRFU: Round 1. Statistical Research Division Study Series Report, Survey Methodology #2005-07.

Hunter, J. E., and A. D. Landreth. (2005). Behavior coding analysis report: Evaluating bilingual versions of the non-response follow-up (NRFU) for the 2004 Census Test. Statistical Research Division Study Series Report, Survey Methodology #2006-07. Available at http://www.census.gov/srd/papers/pdf/ssm2006-07.pdf

Ice, G. H. (2004). Technological advances in observational data collection: The advantages and limitations of computer-assisted data collection. Field Methods 16 3: 352–75.

Johnson, A. J., and Horne, D. (2005). CAPI at MORI – A decade of mobile interviewing. Paper presented at the Association for Survey Computing Conference on Mobile Computing, April 22.

Jones, J. C., and Childs, J. H. (2006). Final Report on Round 2 of the Spanish 2006 NRFU Cognitive Test. Internal report.

Love, K., and Byrne, R. (2005). Investigating Measurement Errors Due to Interviewer Misclassification and Mode. U.S. Census Bureau, DSSD 2010 Census Coverage Measurement Memorandum Series #2006-D7-03, Unpublished internal report, May 23, 2005.

Martin, E., Childs, J.H., DeMaio, T., Hill, J., Reiser, C., Gerber, G., Styles, K., and Dillman, D. (2007). Guidelines for Designing Questionnaires for Administration in Different Modes U.S. Census Bureau, Washington, DC 20233.

Moore, J. C., and Moyer, L. (2002). ACS/CATI Person-Based/Topic-Based Field Experiment. Statistical Research Division Research Report Series, #2002-04.

National Assessment of Adult Literacy. (2006). A First Look at the Literacy of America's Adults in the 21st Century. U.S. Department of Education, Institute of Education Science, 2006. (NCES 2006-470) National Center for Education Statistics.

National Research Council (2006). Once, Only Once, and in the Right Place: Residence Rules in the Decennial Census. Panel on Residence Rules in the Decennial Census. Cork, D. L., and Voss, P. R., eds. Committee on National Statistics, Divison of Behavior and Social Sciences and Education. Washington, DC: The National Academies Press.

Nicholls, W. L., and de Leeuw, E. (1996). Factors in acceptance of computer-assisted interviewing methods: A conceptual and historic review. 1996 <u>Proceedings of the Section on Survey Research Methods, American Statistical Association</u>, Alexandria, VA: American Statistical Association, 758-63.

Nichols, E., and Childs, J.H. (2007). 2006 Questionnaire Design and Experimental Research Survey: Demographic Questions Analysis. SRD Research Report Series Survey Methodology RSM2008/01. Available at <u>http://www.census.gov/srd/papers/pdf/rsm2008-01.pdf</u>

Nichols, E., and Childs, J.H. (forthcoming). Respondent Debriefings Conducted by Experts: A Technique for Questionnaire Evaluation. <u>Field Methods</u>.

Nichols, E., Childs, J.H., and Rodriguez, R. (2007). 2006 Questionnaire Design and Experimental Research Survey: Demographic Questions Analysis.

Norris, S. (2005). 2004 Census Test Evaluation Report #4: Data quality in the use of the hand held computer for nonresponse followup enumeration during the 2004 census test. Unpublished report, September 14, 2005.

Nusser, S. M. (1999). Mobile CASIC and the survey process. <u>Proceedings of the Third International Association for Survey Computing Conference</u>, 209–18. Edinburgh, Scotland: Association for Survey Computing.

Nusser, S. M., Thompson, D. M., and DeLozier, G. S. (1996). Using personal digital assistants to collect survey data. 1996 <u>Proceedings of the Section on Survey Research Methods, American Statistical Association</u>, Alexandria, VA: American Statistical Association, 780–85.

Oksenberg, L., Cannell, C., and Kalton, G. (1991). New strategies for pretesting survey question. <u>Journal of Official Statistics</u> 7 (3): 349–94.

Olmsted, E. and Hourcade, J. P. (2005). NRFU Round II. Quick Report submitted to DMD August 25, 2005.

Olmsted, E., Hourcade, J. P., and Abdalla, A. (2005). NRFU 2006 Quick Report Round I. Submitted to DMD August 8, 2005.

Rappaport, M., Davis, D., and Allen, S. (2006) Final Report on an Observational Study of Census Non-response Follow-up Interviews during the 2006 Census Test in Travis Conty, Texas Spring and Summer 2006. Report submitted by Development Associates to the Census Bureau, December 4, 2006.

Spencer, G., and Perkins, Jr., R. C. (1998). The Effect of Different Question Formats on Age and Birthdate Statistics From the 1996 National Content Survey. Internal Census Bureau report.

Tourangeau, R., Rips, L. J., and Rasinski, K. (2000). <u>The Psychology of Survey Response</u>. Cambridge UP: Cambridge, UK.

Waite, P.J. (2007). The Planned Census for 2010: Expansion of Multi-Mode Data Collection. Paper presented at the Innovative Methodologies for Censuses in the New Millennium Conference. Southampton, England. August 31, 2007.

Willis, G. (2005). <u>Cognitive Interviewing: A Tool for Improving Questionnaire Design</u>. Thousand Oaks: Sage.

WHO TO COUNT

- We want to count people <u>where they usually live and sleep</u>.

- For people with more than one place to live, this is the place where they sleep most of the time.

EXCLUDE these people (They will be counted at the other place):

- College students who live away from this address most of the year
- Armed Forces personnel who live away
- People who, on April 1, 2006, were in a:
  - Nursing home, mental hospital, etc.
  - Jail, prison, detention facility, etc.

INCLUDE these people:

- Babies and children living here, including foster children
- Roommates or boarders
- People staying here on April 1, 2006 who have no other permanent place to live

RESIDENCE RULES

Appendix B: Simplified 2010 NRFU Personal-Visit Script

1. **Did you live or stay at <Address> on April 1, 2010?**

Yes – Go to 3
No - Go to 2

2. **Is there anyone living or staying here now who also stayed here on April 1, 2010?**

Yes
No – *Proxy interview*

3. Is this house a vacation or seasonal home, or does someone usually live here?
__ Vacation, seasonal, held for occasional use
__ Someone usually lives here

4. We need to list people living or staying here on April 1, 2010. We want to list people where they usually live and sleep. For example, college students and armed forces personnel should be listed where they live and sleep most of the time.

If yes to 1:
> **Let's start with you, what is your first name? Middle initial? Last name? Anyone else?**

If no to 1:
> **What is the first name of a person who was living and sleeping here on April 1st? Middle initial? Last name? Anyone else?**

What is the first name of the next person who was living and sleeping here on April 1, 2010? Middle initial? Last name? Anyone else?

5. We do not want to miss any people who might have been staying here on April 1st. Were there any additional people that you didn't mention, for example:

> **Babies?**
> **Foster children?**
> **Any other relatives?**
> **Roommates?**
> **Any other nonrelatives?**
> **How about anyone staying here on April 1st who had no other permanent place to live?**
>
> Yes **– What is that person's first name? Middle initial? Last name? Anyone else?**
> No **-** Continue

6a. (Now thinking of all the people you just mentioned,) in April, (Were you/ was anyone) living in college housing?

Yes – if more than one person in household – **Who was living in college housing?**
No -

b. In April, (Were you/ Was anyone) living away for the military?
Yes – if more than one person in household – **Who was living away for the military?**
No

c. On April 1, 2010, (were you/ was anyone) in a place like a nursing home or a jail or prison?
Yes – **Who was living away in a place like a nursing home or jail or prison?**
No

7. **Do you or does someone in this household own this house with a mortgage or loan (including home equity loans), own it free and clear, rent it or occupy it without having to pay rent?**

8. Of the people who live here, who (owns/rents) this house?
The person selected is the Reference Person.

The remaining questions are asked for each person in a topic-based manner. The example will just provide wording for a single person.

9**. Please look at List A on the handout I gave you at the beginning of the interview. How are you related to [REFERENCE PERSON's NAME]?**

Husband or wife	Roomer or boarder
Biological son or daughter	Housemate or roommate
Adopted son or daughter	Unmarried partner
Stepson or stepdaughter	Other nonrelative
Brother or sister	
Father or mother	
Grandchild	
Parent-in-law	
Son-in-law or daughter-in-law	
Other relative	

10. Are you male or female?

11. What is your date of birth?
    b. *If date of birth is unknown-* **What was your age on April 1, 2010?**

12. (For the census, we need to record age as of April 1, 2010.) So, just to confirm – you were AGE on April 1, 2010?

13a. **Please look at List B. Are you of Hispanic, Latino or Spanish origin?**
Yes - go to 13b
No - go to 14

13b. **Are you Mexican, Mexican American, or Chicano; Puerto Rican; Cuban; or of another Hispanic, Latino or Spanish origin; for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on?**

14a. Please look at List C and choose one or more races. (For this census, Hispanic origins are not races.) Are you White; Black, African American, or Negro; American Indian or Alaska Native; Asian; Native Hawaiian or other Pacific Islander; or Some other race?

   White
   Black/African American/Negro
   American Indian or Alaska Native B *Ask 14b*
   Asian B *Ask 14c*
   Pacific Islander B *Ask 14e*
   Some other race B *Ask 14g*

14b. *(If American Indian or Alaska Native)* **You may list one or more tribes. What is your enrolled or principal tribe?**

_____

14c. *(If Asian)* **You may choose one or more Asian groups. Are you Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese or another Asian group, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on?**

Asian Indian
Chinese
Filipino
Japanese
Korean
Vietnamese
Other Asian (For example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.) *Ask 14d*

    14d. *(If Other Asian)* **What is that other Asian group?** (For example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.)

    _____

14e. *(If Native Hawaiian or Other Pacific Islander)* **You may choose one or more Pacific Islander groups. Are you Native Hawaiian, Guamanian or Chamorro, Samoan, or another Pacific Island group, for example, Fijian, Tongan, and so on?**

Native Hawaiian
Samoan
Guamanian or Chamorro
Other Pacific Islander (For example, Fijian, Tongan, and so on.) *Ask 14f*

14f. *(If some other Pacific Island group)* **What is that other Pacific Islander group?** (For example, Fijian, Tongan, and so on.)

_____

14g. *(If Some Other Race)* **What is your other race group?**

_____

15. **Just to make sure everyone is counted in the right place, did you sometimes live or stay somewhere else such as at a seasonal or second residence, for child custody, or for any other reason?**