

Census 2000 Synthesis Report No.17

*Census 2000 Testing, Experimentation,
and Evaluation Program*

Issued March 2004

TR-17

Census 2000 Alternative Questionnaire Experiment

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

United States
**Census
2000**

Acknowledgments

The Census 2000 Evaluations Executive Steering Committee provided oversight for the Census 2000 Testing, Experimentation, and Evaluations (TXE) Program. Members included **Cynthia Z. F. Clark**, Associate Director for Methodology and Standards; **Preston J. Waite**, Associate Director for Decennial Census; **Carol M. Van Horn**, Chief of Staff; **Teresa Angueira**, Chief of the Decennial Management Division; **Robert E. Fay III**, Senior Mathematical Statistician; **Howard R. Hogan**, (former) Chief of the Decennial Statistical Studies Division; **Ruth Ann Killion**, Chief of the Planning, Research and Evaluation Division; **Susan M. Miskura**, (former) Chief of the Decennial Management Division; **Rajendra P. Singh**, Chief of the Decennial Statistical Studies Division; **Elizabeth Ann Martin**, Senior Survey Methodologist; **Alan R. Tupek**, Chief of the Demographic Statistical Methods Division; **Deborah E. Bolton**, Assistant Division Chief for Program Coordination of the Planning, Research and Evaluation Division; **Jon R. Clark**, Assistant Division Chief for Census Design of the Decennial Statistical Studies Division; **David L. Hubble**, (former) Assistant Division Chief for Evaluations of the Planning, Research and Evaluation Division; **Fay F. Nash**, (former) Assistant Division Chief for Statistical Design/Special Census Programs of the Decennial Management Division; **James B. Treat**, Assistant Division Chief for Evaluations of the Planning, Research and Evaluation Division; and **Violeta Vazquez** of the Decennial Management Division.

As an integral part of the Census 2000 TXE Program, the Evaluations Executive Steering Committee chartered a team to develop and administer the Census 2000 Quality Assurance Process for reports. Past and present members of this team include: **Deborah E. Bolton**, Assistant Division Chief for Program Coordination of the Planning, Research and Evaluation Division; **Jon R. Clark**, Assistant Division Chief for Census Design of the Decennial Statistical Studies Division; **David L. Hubble**, (former) Assistant Division Chief for Evaluations and **James B. Treat**, Assistant Division Chief for Evaluations of the Planning, Research and Evaluation Division; **Florence H. Abramson**, **Linda S. Brudvig**, **Jason D. Machowski**, and **Randall J. Neugebauer** of the Planning, Research and Evaluation Division; **Violeta Vazquez** of the Decennial Management Division; and **Frank A. Vitrano** (formerly) of the Planning, Research and Evaluation Division.

The Census 2000 TXE Program was coordinated by the Planning, Research and Evaluation Division: **Ruth Ann Killion**, Division Chief; **Deborah E. Bolton**, Assistant Division Chief; and **Randall J. Neugebauer** and **George Francis Train III**, Staff Group Leaders. **Keith A. Bennett**, **Linda S. Brudvig**, **Kathleen Hays Guevara**, **Christine Louise Hough**, **Jason D. Machowski**, **Monica Parrott Jones**, **Joyce A. Price**, **Tammie M. Shanks**, **Kevin A. Shaw**,

George A. Sledge, **Mary Ann Sykes**, and **Cassandra H. Thomas** provided coordination support. **Florence H. Abramson** provided editorial review.

This report was prepared by **Elizabeth Ann Martin** of the Directorate for Methodology and Standards and **Eleanor Gerber** and **Cleo Redline** (formerly) of the Statistical Research Division. The following authors and project managers prepared Census 2000 experiments and evaluations that contributed to this report:

Demographic Statistical Methods Division

Sharon R. Ennis
Phyllis Singer

Directorate for Methodology and Standards

Elizabeth Ann Martin

Statistical Research Division

Aref N. Dajani
Eleanor Gerber
Cleo Redline
Mary Ann Scaggs

Independent Contractor

Don Dillman

The authors would like to recognize **Aref N. Dajani** and **Mary Ann Scaggs**, who prepared files, calculated response rates, and assisted in analyses for the Census 2000 Alternative Questionnaire Experiment. Thanks to **Don Dillman** for his comments on this report, and his never-flagging interest in, and enthusiasm for, questionnaire design issues.

Greg Carroll and **Everett L. Dove** of the Administrative and Customer Services Division, and **Walter C. Odom**, Chief, provided publications and printing management, graphic design and composition, and editorial review for print and electronic media. General direction and production management were provided by **James R. Clark**, Assistant Division Chief, and **Susan L. Rappa**, Chief, Publications Services Branch.

Census 2000 Synthesis Report No. 17

*Census 2000 Testing, Experimentation,
and Evaluation Program*

Issued March 2004

TR-17

CENSUS 2000 ALTERNATIVE QUESTIONNAIRE EXPERIMENT



U.S. Department of Commerce
Donald L. Evans,
Secretary

Vacant,
Deputy Secretary

Economics and Statistics Administration
Kathleen B. Cooper,
Under Secretary for Economic Affairs

U.S. CENSUS BUREAU
Charles Louis Kincannon,
Director

Suggested Citation

Elizabeth Ann Martin,
Eleanor Gerber, and Cleo Redline
*Census 2000 Testing,
Experimentation, and Evaluation
Program Synthesis Report No. 17,
TR-17, Census 2000 Alternative
Questionnaire Experiment.*
U. S. Census Bureau
Washington, DC 20233



Economics and Statistics Administration

Kathleen B. Cooper,
Under Secretary for Economic Affairs



U.S. CENSUS BUREAU

Charles Louis Kincannon,
Director

Hermann Habermann,
Deputy Director and Chief Operating Officer

Cynthia Z. F. Clark,
Associate Director for Methodology and Standards

Preston J. Waite,
Associate Director for Decennial Census

Teresa Angueira,
Chief, Decennial Management Division

Ruth Ann Killion,
Chief, Planning, Research and Evaluation Division

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll-free 866-512-1800; DC area 202-512-1800
Fax: 202-512-2250 Mail: Stop SSOP, Washington, DC 20402-0001

Contents

Foreword	v
Executive Summary	1
1. Introduction	3
1.1. Experiment A: Effects of Altering the Design of Branching Instructions on Navigational Performance in Census 2000, by Cleo Redline, Don Dillman, Aref Dajani, and Mary Ann Scaggs	3
1.2. Experiment B: An Experiment to Improve Coverage Through Revised Roster Instructions, by Eleanor Gerber, Aref Dajani, and Mary Ann Scaggs	3
1.3. Experiment C: Questionnaire Effects on Reporting of Race and Hispanic Origin: Results of a Replication of the 1990 Mail Short Form in Census 2000, by Elizabeth Martin . . .	4
2. Methods	5
2.1. Questionnaire development	5
2.2. Sample design	5
2.3. Experimental treatments	5
2.4. Additional processing and sources of data	6
3. Major Findings	7
3.1. Effects of Altering the Design of Branching Instructions on Navigational Performance	7
3.2. An Experiment to Improve Coverage Through Revised Roster Instructions	8
3.3. Questionnaire Effects on Reporting of Race and Hispanic Origin	9
4. Implications and Recommendations	15
References	19
Appendix	21

LIST OF TABLES, CHARTS, AND FIGURES

Chart 1. Alternative Questionnaire Experiment Panels6

Table 1. Error Rates For All Census Long Form Items With
Branching Instructions7

Table 2. Omission Rates for Hispanics by Panel and Strata8

Table 3. Frequency of Reasons Given for Erroneous Enumerations
of People Age 18 to 359

Table 4. Reporting of Hispanic Origin in Mail Questionnaires by
Form Type9

Table 5. Race Item Nonresponse Rates by Form Type and
Hispanic Origin10

Table 6. Detailed Hispanic Reporting by Form Type10

Table 7. Race Distributions for Hispanics by Form Type11

Table 8. Percentage Who Reported in a Race "Example Group" by
Form Type11

Table 9. Percent Writing in Each "Example Group" by Form Type 12

Figure 1. Illustration of the Five Branching Instruction Treatments 21

Figure 2. Census 2000 Roster Instructions (Experiment B)22

Figure 3. Roster Instructions in Experimental Form
(Experiment B)23

Figure 4. Census 2000-style Form: Race and Hispanic Questions
(Experiment C)24

Figure 5. 1990-style Form: Race and Hispanic
Questions (Experiment C)25

Foreword

The Census 2000 Testing, Experimentation, and Evaluation Program provides measures of effectiveness for the Census 2000 design, operations, systems, and processes and provides information on the value of new or different methodologies. By providing measures of how well Census 2000 was conducted, this program fully supports the Census Bureau's strategy to integrate the 2010 planning process with ongoing Master Address File/TIGER enhancements and the American Community Survey. The purpose of the report that follows is to integrate findings and provide context and background for interpretation of related Census 2000 evaluations, experiments, and other assessments to make recommendations for planning the 2010 Census. Census 2000 Testing, Experimentation, and Evaluation reports are available on the Census Bureau's Internet site at: www.census.gov/pred/www/.

This page intentionally left blank.

Executive Summary

This report summarizes the findings of three experiments included in the Census 2000 Alternative Questionnaire Experiment. The purposes of these experiments were diverse:

- **The skip instruction experiment** examined respondent performance in following skip instructions in the Census long form. It compared different ways of aiding respondents' navigation through the questionnaire. One design incorporated instructions and visual features to help respondents prevent errors before they occurred, and another was designed to help respondents detect errors after they occurred. In addition to these prevention and detection designs, other potential design improvements, such as using reverse print to attract respondents' attention to instructions, and rewording the standard "skip" instruction, were also tested.
 - **The residence instructions experiment** involved the presentation of residence rules in the Census 2000 short form. This research aimed to improve within-household coverage by rewording the roster instructions to be understandable to respondents, encouraging respondents to read them through appropriate placement and formatting, and by presenting the instructions so as to increase respondents' willingness to follow them.
 - **The race and Hispanic origin experiment** compared the 1990-style race and Hispanic origin questions with the new questions in Census 2000 short form. It examined the effects of changes mandated by the Office of Management and Budget to allow the reporting of more than one race and reverse the sequence of the race and Hispanic origin items. Other changes in format, categories and wording were also introduced in Census 2000, and the net effects of all the changes on race and Hispanic reporting were analyzed.
- All three experiments were limited to the mailout-mailback universe.
- The results of the three experiments include the following:
- **Skip instruction experiment:** Errors of commission (which occur when respondents incorrectly answer questions they should have skipped) were significantly reduced in all of the experimental treatments, suggesting that the design changes improved respondents' perception and comprehension of the instruction. Errors of omission (which occur when respondents skip questions they should have answered) decreased for the Detection Treatment, but significantly increased for every other experimental treatment. Either type of error indicates respondent difficulty navigating the questionnaire, but their impact is different. Errors of omission result in missing data. Errors of commission increase respondent burden and frustration. **We recommend that the Census Bureau adopt the Detection method in mail questionnaires, since it significantly reduces both types of errors.**
 - **Residence instructions experiment.** The changes in format, presentation, and wording of the residence instructions resulted in a significantly higher response to the household count question (which serves as an important indicator of missing data and flags large household follow-up.) The experimental panel also produced significantly fewer omissions among Hispanics in the low coverage stratum. **We recommend additional testing of the graphical and wording changes that led to these improvements, to better understand their effects and to further improve the quality of household count data.**
 - **Race and Hispanic origin experiment.** Overall, the questionnaire revisions substantially improved the completeness of race and Hispanic origin reporting in mail short form questionnaires. In addition, Hispanics were less likely to report their race as Some other race, and more likely to report as White, in the 2000-style questionnaires.
- Although there were no apparent questionnaire effects on the fraction reported as Hispanic,

there were effects on the reporting of detailed Hispanic origin groups. The 1990-style questionnaire obtained more detailed reports of Hispanic origin than the 2000-style questionnaire, probably due to the effects of question wording differences as well as examples.

Unexpectedly, there were three times as many reports of the example groups for Asian and Pacific Islander groups in the 2000-style form, which did not list examples, as there were in the 1990-style form, where examples were listed. The experiment demonstrates that

some questionnaire design changes made in Census 2000 resulted in substantial improvements in data quality, but that other changes had unintended consequences. **We recommend careful pretesting and field testing of all changes to the questionnaire, and that similar but larger replication studies be conducted in future censuses to evaluate the effects of questionnaire changes on the comparability of data from one census to the next.**

The results of all three experiments point to interactions

between question format and content, suggesting that we must attend to the complex relationships between format and meaning in self-administered questionnaires. These factors have been demonstrated here to have measurable effects on the data. These experiments demonstrate that format affects performance on branching instructions, affects response/nonresponse on the household count question and indirectly affects coverage and that format differences between 1990-style and Census 2000 forms affect race and ethnicity reporting.

1. Introduction

An Alternative Questionnaire Experiment (AQE) to test the effects of variations in the design of census mail questionnaires upon response rates and data quality has been conducted in each of the past three decennial censuses. The 1980 Alternative Questionnaire Experiment tested two FOSDIC (Film Optical Sensing Device for Input to Computer) matrix style forms, and a non-FOSDIC form that was intended to be attractive and easily understood (Mockovak 1984). The 1990 AQE experimentally compared five long form questionnaires that involved successively more radical departures from the traditional design, including wording and format changes, a kit containing individual questionnaires for each household member, and anonymous census questionnaires (Bates 1991, 1992).

The 2000 AQE incorporates three separate experiments, one involving census long forms and the other two involving short forms, with different objectives. Two experiments look forward, seeking improvements in the design of mail forms that may lead to improved data quality in the next census. A third looks backward, replicating the 1990 short form in Census 2000 in order to document the effects on the data of changes in the design of the Census 2000 short form.

All three experiments test combinations or "packages" of design features, rather than testing each design change separately in a controlled fashion that would permit inferences about their individual

effects. Thus, firm conclusions can only be drawn about the combined effect of multiple design features, and this is an important limitation of all three experiments. However, previous research and testing often provides insights into the effects of particular design features. Another limitation is that the findings are only generalizable to the mailout/mailback universe. This excludes certain populations of interest, such as Indians living on reservations and Alaska Natives (of interest to the analysis of questionnaire effects on race reporting) or rural populations not enumerated by mail (of interest to the analysis of roster completeness).

The three experiments are:

1.1 Experiment A: *Effects of Altering the Design of Branching Instructions on Navigational Performance in Census 2000*, by Cleo Redline, Don Dillman, Aref Dajani, and Mary Ann Scaggs

This experiment took as its starting point the difficulty many respondents have navigating the census long form, causing them to mistakenly skip questions they are supposed to answer or answer questions they are supposed to skip. Hypotheses derived from research on visual perception were applied to develop new strategies for helping respondents navigate their way through the questionnaire. One design incorporated instructions and visual features to help respondents prevent errors before they occurred, and another

was designed to help respondents detect errors after they occurred. In addition to these *prevention* and *detection* designs, other potential design improvements, such as using reverse print to attract respondents' attention to instructions, and rewording the standard "skip" instruction to clarify it, were also tested. The larger purpose of this experiment was to develop a better understanding and general principles of how graphical design features of a questionnaire influence respondents' ability to navigate through it.

1.2 Experiment B: *An Experiment to Improve Coverage Through Revised Roster Instructions*, by Eleanor Gerber, Aref Dajani, and Mary Ann Scaggs

This experiment focused on the problems respondents have filling out household rosters correctly. Erroneously including a person who does not live in a household, or omitting one who does, result in census coverage errors. The problem is exacerbated by counter-intuitive and complex census residence rules, which are often ignored by respondents who may have their own firm ideas about who belongs in their household. In addition, census residence rules do not follow any simple logic which is easily communicated to respondents. This research aimed to improve within-household coverage by improving the roster instructions in three ways: first, by rewording them to be understandable to respondents, second by encouraging respondents to read them

through appropriate placement and formatting, and third, by presenting the instructions so as to increase respondents' willingness to follow them.

**1.3 Experiment C:
*Questionnaire Effects on Reporting of Race and Hispanic Origin: Results of a Replication of the 1990 Mail Short Form in Census 2000, by Elizabeth Martin***

This experiment replicated a 1990-style mail short form during

Census 2000 and compared the results to data from Census 2000 short forms in order to evaluate how the questionnaire changes affected reporting of race and Hispanic origin reporting. The questionnaire changes introduced in Census 2000 included allowing the reporting of more than one race and reversing the sequence of the race and Hispanic origin items, as well as other changes in format, categories and wording.

This synthesis report summarizes the results of the three experiments, and seeks to draw more general conclusions from them. Section 2 describes the methods used in the AQE, and section 3 summarizes the major findings of each separate experiment (as well as additional analysis of one that did not appear in the final report). More detailed discussions of the methods and findings of each experiment are found in their respective final reports.

2. Methods

2.1 Questionnaire development

Experiment A. Conventional branching (skip) instructions may easily be overlooked. Alternative forms were designed to manipulate verbal (revising "skip to" to "go to") and graphic (reverse print) features, and to implement *error detection* or *error prevention* strategies. Five alternative forms were developed and tested in a classroom experiment and in cognitive tests. The testing was used to revamp and refine the designs which were tested in the AQE. The questions were chosen specifically so that content would not provide cues to navigation. Each panel used identical content, ensuring that the length of each panel was the same. See Fig. 1 for the five formats that were included in the AQE.

Experiment B. Several different roster formats were designed, and two rounds of cognitive testing were conducted to select and refine the best format, which became experimental panel 7 (see chart 1, below) in this experiment. The experimental roster included several design features, including placing the instructions before the roster question, rewording the instruction to read "our guidelines" before answering the question, enclosing the instructions and answer box in a black outlined box with a slightly brighter background than the surrounding questionnaire, double-banking the include/exclude rules and rewording them to make them more inclusive and easier to understand (see

Fig. 2 for the Census 2000 roster question and Fig. 3 for the experimental version).

Experiment C. A 1990-style form was developed which preserved 1990 question wording, categories, order, type size, matrix format, and other essential design features which might influence responses. The 1990-style form was not identical to the 1990 census form, however, because it changed dates and incorporated several publicized and familiar features of the 2000 design (color, logo, "Start here" instruction, envelope and letter). Any questions not included in the Census 2000 short form, such as marital status, were dropped. The form was not pretested. See Figs. 4 and 5 for the two versions of the race and Hispanic questions which were compared.

2.2 Sample design

The AQE included eight panels of approximately 5,000 households (10,000 for one panel) each. The sample was stratified into high coverage areas (HCAs) which had low proportions of minorities and renters in the 1990 census, and low coverage areas (LCAs) with high proportions of minorities and renters. Sample cases were distributed equally between strata, implying that households in LCAs were oversampled. All results are weighted to reflect sampling probabilities.

Addresses on the Decennial Master Address File in the mailout/mail-back areas of the country at the

time sample selection took place served as the universe for sample selection (Woltman, 1999). Addresses in non-mailback areas (mostly rural areas) were excluded from sample. This excludes certain population groups of interest, including American Indians living on reservations and Alaska Natives. Addresses that were added later as a result of coverage improvement operations were excluded, as were addresses in the sample for the Accuracy and Coverage Evaluation survey. A systematic sample by state, stratum, and treatment was selected.

2.3 Experimental treatments

The AQE included eight panels (three short form and five long form panels). Chart 1 shows the experimental and control panels and number of households in each. Letters in parentheses indicate whether the panel is part of Experiment A, B, or C of the AQE.

For all panels, questionnaires were mailed out according to the Census 2000 schedule, with every sampled address mailed an advance letter, a questionnaire, and a follow-up postcard. For respondents in the AQE, the responses provided on the mail forms were their census data. Telephone Questionnaire Assistance operators were trained to answer questions from respondents who received an experimental questionnaire (e.g., about the instruction in the 1990-style form to select one race category from respondents who wanted to report more than one).

Chart 1. Alternative Questionnaire Experiment Panels

Panel	Mailing sample size	Response rate
Long form panels		
1. Census 2000 long form questionnaire (Control panel, experiment A)	5,257	63.95
2. "Go to" questionnaire (A): identical to panel 1 questionnaire, except "skip to" is replaced with "go to" throughout.	5,248	64.34
3. Reverse print questionnaire (A): identical to panel 2 questionnaire, except the "go to" instruction appears in reverse print (yellow letters on black background).	5,251	61.82
4. Prevention questionnaire (A): adds an instruction alerting respondents to look for branching instructions; adds arrows and other features to attract attention to instructions.	5,241	63.13
5. Detection questionnaire (A): adds an arrow to guide respondents away from branching instruction when appropriate; adds information to the next question to allow respondents to judge if they have correctly skipped.	5,238	63.25
Short form panels		
6. Census 2000 short form questionnaire (Control panel, experiments B, C)	5,252	73.07
7. Revised roster questionnaire (B)	5,256	73.52
8. 1990-style short form questionnaire (C): replicates 1990 question wording, categories, order, type size, matrix format, and other essential design features. Incorporates Census 2000 color, logo, envelope and letter. Drops questions not included in the Census 2000 short form (marital status)	10,499	72.60

Questionnaires were mailed back to the National Processing Center in Jeffersonville, Indiana, where they were keyed and processed separately from production Census 2000 data, which were imaged. Households which did not return a mail questionnaire were followed up as part of the Census 2000 nonresponse operation and are not included in this analysis. Response rates for the eight panels are shown in the second column of Chart 1. The rates are weighted and exclude undeliverable addresses, duplicate forms, and blank forms. Response rates for the two experimental short form panels do not significantly differ from the control². As was true in the census, response rates for the long form panels are significantly lower than for the short form panels. There are no significant response

rate differences among the long form panels. (Significances were calculated using Bonferroni adjustments for multiple comparisons; see Dajani and Scaggs, 2001.)

2.4 Additional processing and sources of data

Experiment B. To analyze the effects of roster variations, coverage was measured by a specially developed telephone coverage reinterview, conducted by Westat about four months after the census, in late July of 2000. The reinterview sample consisted of cases

² Although overall short form response rates do not differ, panel 8 has a significantly lower response rate (57.63 percent) than the control panel 6 (60.78 percent) in the low coverage area stratum ($p < .05$). This difference suggests that the Census Bureau's investment in developing a "user-friendly" design for the Census 2000 short form did achieve a higher response rate in these areas, compared to a 1990-style form.

that had completed and returned the census form, had phone numbers, and were not sent to large household follow-up. These cases were subsampled randomly at a rate of 50 percent in the high coverage area stratum. The total sample size for the reinterview was 4,218 households. There were 2,958 completed interviews: 1,497 completed cases in the control and 1,461 in the experiment. This represents a response rate of 70.35 percent in the control and 69.90 percent in the experiment, with an overall response rate of approximately 70 percent.

Experiment C. To increase sample size and improve reliability for the analysis of questionnaire effects on race and Hispanic reporting, the short form control (panel 6) was supplemented with mail returns from the control panel for the Response Mode and Incentives Experiment (RMIE) (Guarino, 2001). The RMIE control group sample of approximately 20,000 addresses was selected from the same universe using the same stratification as AQE, except the sample was allocated proportionately to the HCA and LCA strata. All addresses in the RMIE control group received Census 2000 short form questionnaires, which were processed separately, like the AQE. The response rate was 71.5 percent.

Race and Hispanic origin data for panels 6 and 8 (and the RMIE control data) were coded and pre-edited by applying a simplified version of pre-edits used in Census 2000 production. (See Martin 2002 for details.) Missing data were not imputed or allocated, as they would be in fully edited census data. Results may differ for fully edited census race and Hispanic origin data.

3. Major Findings

3.1 Effects of Altering the Design of Branching Instructions on Navigational Performance

The analysis examined two types of errors. Commission errors occur when respondents incorrectly provide an answer to questions they should have skipped, and omission errors occur when respondents skip questions they should have answered. Either type of error indicates respondent difficulty navigating the questionnaire, but their impact is different. Errors of omission result in missing data. Errors of commission increase respondent burden and frustration.

To control for differences in the number of questions that respondents answered, analysis was limited to Person 1 responses (the questionnaire provides space for data to be provided for up to six household members). Error rates were calculated for questions with branching instructions (because

only their designs differed between form types) and with valid responses (because only then was it evident whether a respondent should branch or not).

The results indicate that simply changing the wording of branching instructions from "skip to" to "go to" (panel 1 versus 2) did not significantly affect either errors of commission or omission. Probably respondents did not notice either instruction, and no amount of rewording will help if the problem is that respondents are not reading the information in the first place.

Panel 2 was adopted as the control for comparison with the remainder of the experimental panels, since all shared the "go to" instruction wording. All three experimental treatments resulted in significantly fewer errors of commission than panel 2. The reverse print instruction (panel 3) had significantly fewer errors than the instruction without reverse print (panel 2).

The prevention questionnaire had significantly fewer errors than the reverse print instruction. There was no further significant reduction in commission errors with the detection instruction overall (although there was in the LCA stratum). The fact that the commission error rate decreases across the Go To, Reverse Print, Prevention and Detection Treatments suggests that the changes made from one design to the next improved respondents' perception and comprehension of the instruction.

A different pattern occurs for errors of omission. While errors of omission decreased for the Detection Treatment, they significantly increased for every other experimental treatment. Possibly, the attention-getting features of the experimental treatment distracted respondents' attention from other cues about which questions they were supposed to answer. The better performance of the detection method may be due to two of its features. First, it included a visual cue (arrow) to direct respondents to the next question. Second, the feedback mechanism (the left-hand arrow that came off the non-branching response option and pointed to a parenthetical feedback phrase) may have helped respondents avoid errors of omission.

Thus, commission errors were decreased in this treatment without the omission errors increasing. The level of omissions in the Detection treatment was significantly less than in the "Go to"

Table 1.
Error Rates for all Census Long-Form Items With Branching Instructions

Instruction Treatment	Errors of commission	Errors of omission
1. Census 2000 (Skip to instruction)	19.7	5.0
2. (Go to) Control	20.8	5.4
3. (Go to) Reverse Print	17.9	7.6
4. (Go to) Prevention	14.7	7.0
5. (Go to) Detection	13.5	4.0
Statistical Comparison		
1 vs. 2	n.s.	n.s.
2 vs. 3	p<.01	p<.01
2 vs. 4	p<.01	p<.01
2 vs. 5	p<.01	p<.01
3 vs. 4	p<.01	n.s.
3 vs. 5	p<.01	p<.01
4 vs. 5	n.s.	p<.01

Note: Lack of statistical significance is denoted by n.s.

control. Interestingly, some cognitive respondents and respondents in a debriefing expressed the opinion that the many arrows in the detection method were "confusing"—even though the results suggest these arrows improved their performance! This indicates a need for care about what evidence should be used in making decisions, and also indicates that perhaps the effects of the arrows need to be tested separately.

Wide variations existed in error rates for individual questions. The reasons were not explored in this research. However, the results suggest that respondents do not understand the questions or the underlying response task (that they need not answer every question). As a result the respondent burden is greater than necessary (on the average, respondents are answering 20 percent more questions than they need to). Further research is needed to improve respondents' understanding of the questions and the response task to reduce respondent burden.

Thus, manipulating the verbal symbolic and graphic languages of branching instructions significantly influences whether the instructions are followed. Further research is needed into the ways that graphic features interact with reading comprehension—that is, what respondents read, the order in which they read it and their consequent interpretation of that they read. This is an area of questionnaire design that is clearly emerging as critical to data quality and in need of further research.

3.2 An Experiment to Improve Coverage Through Revised Roster Instructions

The revised roster design was evaluated using several error measures: nonresponse for the house-

hold count box, where respondents recorded the number of persons in their household; omission rates; and erroneous enumeration rates. Omission rates were calculated as the number of Census Day residents omitted from the census roster (but identified during reinterview) divided by the number of correct enumerations, using information both from the census form and from the reinterview. Erroneous enumeration rates were calculated as the number of persons on the census roster who were identified as Census Day nonresidents in reinterview, divided by the number of people on the census roster. In addition, the demographic characteristics of people enumerated using the redesigned roster and Census 2000 roster design were compared.

Item Nonresponse. One of the alterations in the experimental version of the form was the placement of the box where respondents were to record the number of persons in their households. It is critical that the box where respondents record the number of persons in their households be completed, since it serves to flag missing person-level data and to cue large household follow-up. Any increase in item nonresponse in this item would be unacceptable. The item nonresponse for this item is significantly lower in the experimental form (.80 percent) than in the control (1.78 percent) at the

0.01 level of significance. Thus, the item nonresponse rate was cut in half, and the difference was significant in both strata. This demonstrates that the wording changes and/or the format integrating the instructions with the first question were effective in getting respondents to fill in the box.

Omissions are persons who should have been listed on the census but were not. Such persons were identified only in the reinterview (since by definition, they were not on the Census form.)

There is no significant form difference in the omission rates, which were 1.13 percent for the control and 1.21 percent for the experimental form. However, the experimental form had a significantly lower omission rate for Hispanics in the low coverage stratum, as shown in Table 2.

Erroneous enumerations are persons who were included on the census forms although they were not legitimate census day residents. They included persons who spent most of their time elsewhere, or who were in group quarters where they should have been counted on Census Day. Examples are college students living away from home and persons in the military stationed elsewhere.

No significant panel differences were found in erroneous enumeration rates, which were 0.40

Table 2.
Omission Rates for Hispanics by Panel and Strata

Panel	Stratum		
	Total Percent	HCA Percent	LCA Percent
Control	3.54	3.90	3.23
Experiment	2.55	4.26	1.00
Total	3.00	4.09	2.02
Control vs. Experiment: Statistical Comparison	Not Significant	Not Significant	p<.05

percent for the control and 0.39 percent for the experimental form.

Although the rates did not differ by panel, the characteristics of the people who were erroneously enumerated did, for the low coverage stratum: in the experimental panel, the fraction who were young people 18-35 years old was 58 percent, compared to 30 percent in the control panel.

The reasons why the 57 people age 18 to 35 were erroneously enumerated were further examined. Table 3 shows that college students account for a substantial portion of the erroneous enumerations in both forms (and in both the LCA and HCA strata; this result is not shown). However, the largest number of erroneous enumerations is not accounted for by the categories of erroneous enumerations which are usually included in residence rules research, such as college, military, and various group quarters institutions. This suggests that highly mobile people account for a substantial

proportion of erroneous enumerations. This confirms previously conducted qualitative research that has demonstrated that young adults in this age group may be highly mobile for reasons other than college.

Conclusions. The results suggest that the revisions were effective. We do not know which of the design and wording changes are responsible for the improvements. Further research is needed to better understand how the different design features affect responses, how respondents naturally read questionnaires, and the relationship between graphical presentation and meaning in them. We recommend additional research that builds on these encouraging findings in further cognitive and field tests of wording and graphical design changes in the roster questions and instructions.

Differences between the effectiveness of these techniques in high and low coverage areas are encouraging, since coverage improve-

ments occurred for a group (Hispanics) with relatively high rates of omission in the census. The differences also suggest that demographic, cultural, or social factors may influence the effectiveness of one or more of the graphical and wording changes. For instance, perhaps the new format was easier to understand and follow by respondents who speak Spanish rather than English as their native language, or by those who are less educated. The experiment does not allow us to identify the factors which may have played a role, but the results suggest that further research is needed to expand our understanding of their influence.

3.3. Questionnaire Effects on Reporting of Race and Hispanic Origin

Item nonresponse rates and differences in reporting of Hispanic origin and race were compared between the 1990- and 2000-style questionnaires and assessed using VPLX (Fay, 1998).

Item Nonresponse Rates. Overall, the questionnaire revisions substantially improved the completeness of race and Hispanic origin reporting in mail short form questionnaires. Item nonresponse (i.e. blank or uncodable responses) for Hispanic origin was 3.33 percent in 2000-style questionnaires, compared with 14.46 percent in 1990-style questionnaires, as seen in Table 4.

Item nonresponse for race was 3.27 percent for 2000-style forms and 5.95 percent for 1990-style questionnaires, as shown in Table 5. For Hispanics the reduction in race item nonresponse was very large, from 30.53 to 20.79 percent in the 2000-style questionnaires, but remained quite high.

Table 3.
Frequency of Reasons Given for Erroneous Enumerations of People Age 18 to 35

	Total	Control	Experiment
College	23	9	14
Military	4	3	1
Institution (GQ)	2	0	2
Unknown	28	12	16
Total	57	24	33

Table 4.
Reporting of Hispanic Origin in Mail Questionnaires by Form Type

	Form type		t ₂₀₀₀₋₁₉₉₀
	2000-style	1990-style	
Total Percent	100.00	100.00	
Hispanic	11.17	11.14	.05
Non-Hispanic	85.50	74.39	*15.8
Hispanic item blank, uncodable	3.33	14.46	*-21.9
N.	40,723	16,616	

*p<.05

Table 5.
Race Item Nonresponse Rates by Form Type and Hispanic Origin

Hispanic origin	Percent of people missing data on race		
	2000-style	1990-style	t ₂₀₀₀₋₁₉₉₀
Total	3.27	5.95	*-7.34
Hispanics	20.79	30.53	*-4.42
Non-Hispanics	.60	1.53	*-5.03
Hispanic origin missing	13.18	9.72	*2.00

*p<.05

Table 6.
Detailed Hispanic Reporting by Form Type

	2000-style	1990-style	t ₂₀₀₀₋₁₉₉₀
Percent of all people identified as Hispanic	100.00	100.00	
Check box groups: Hispanic groups with separate check boxes in both forms	70.25	73.23	-1.37
Example groups: listed as examples in 1990-style form but not Census 2000	6.41	11.16	*-3.58
All other detailed Hispanic groups	4.20	8.68	*-3.38
Write-in is general descriptor (Hispanic / Latino / Spanish)	11.90	1.90	*10.32
Hispanic, no write-in (or write-in uncodable)	7.25	5.03	*2.15
N	5,163	3,091	

*p<.05

Hispanic Origin Reporting. Results in Table 4 show that the same percentage (slightly over 11.1 percent) was reported as Hispanic in both 1990- and 2000-style forms, despite the reversed sequence of Hispanic origin and race and question wording differences. In past censuses, most people for whom origin was missing were non-Hispanic. On this assumption, these results suggest that questionnaire changes reduced item nonresponse but did not otherwise affect reporting as Hispanic.

However, the assumption may no longer hold with the new question. In 1990, Hispanics and non-Hispanics were equally likely to leave the Hispanic item blank (McKenney et al., 1993), implying that Hispanics were equally represented among people who responded and those who did not respond to the item in 1990. Recently available evidence from the Content Reinterview Survey for

Census 2000 suggests that 25 percent of people who did not respond to the Hispanic item in Census 2000 reported as Hispanic in the reinterview (Singer and Ennis, 2003). This suggests that Hispanics were overrepresented among nonrespondents to the Hispanic item in this census. The high nonresponse rate (esp. for 1990-style forms), uncertainty about what fraction of nonrespondents to the item are Hispanic, and the possibility that the fraction varies by form type, creates uncertainty about whether there might be differences between forms in the fraction identified as Hispanic, after the data were fully edited and imputed. Any such differences are not measurable using this sample.

Effects of Examples: Detailed Hispanic Reporting. Although there were no apparent questionnaire effects on the fraction reported as Hispanic, there were effects

on the reporting of detailed Hispanic origin groups.

It has been suggested that dropping examples from the Hispanic origin question in the Census 2000 mail form may have resulted in a loss of detail in Hispanic reporting. To examine this possibility, Hispanic origins were classified into four categories, as shown in Table 6:

- Groups with *check boxes* (Mexican, Puerto Rican, Cuban) in both 1990- and 2000-style forms;
- Groups listed as *examples* in the 1990- but not the 2000-style form (Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard);
- *All other specific groups* with no check boxes and not listed as examples in either form; and
- *General descriptors*: "Hispanic," "Latino," or "Spanish" were written in, rather than a specific group.

The 1990-style form elicited more detailed reports of Hispanic origin than the 2000-style questionnaire for all three categories: Hispanic groups with separate check boxes, those listed as examples in 1990 but not 2000, and the remaining detailed groups. The differences are significant for the latter two categories. Overall, about 93 percent of Hispanics reported a specific group in 1990-style forms, compared with 81 percent who filled out 2000-style forms. In the 2000-style forms, Hispanics tended to describe their ethnicity in general rather than specific terms. About 12 percent gave Hispanic, Latino, or Spanish as their "group," compared with 2 percent in 1990-style forms.

Table 7.
Race Distributions for Hispanics by Form Type

	Form type		t-statistic
	2000-style	1990-style	t ₂₀₀₀₋₁₉₉₀
Percent of all people identified as Hispanic . . .	100.00	100.00	
White	48.98	39.88	*3.23
Black	2.07	2.32	-.34
American Indian and Alaska Native	1.48	.72	1.61
Asian58	.88	-.60
Native Hawaiian and Other Pacific Islander01	.15	-1.14
Some other race	39.03	51.47	*-4.32
Two or more races	7.84	4.59	*2.88

*p<.05

Table 8.
Percentage Who Reported in a Race "Example Group" by Form Type

	1990-style	2000-style
Wrote in Hmong, Fijian, Laotian, Thai, Tongan, Pakistani, or Cambodian106 (.0351)	.356 (.0606)
Checked or wrote in another race	99.894	99.644
Total percent.	100.000	100.000

These results suggest that the examples helped respondents understand that the intent of the question was to elicit a detailed Hispanic origin, and thereby improved reporting of both example groups and non-example groups. However, the reporting differences are likely due to the effects of question wording differences as well as examples. There was a significant difference in reporting for one of the checkbox categories, the wording of which was identical in both forms ("Yes, Mexican, Mexican-Am., Chicano"). About 54 percent of Hispanics checked the Mexican box (or wrote in Mexican) in 2000-style forms, compared to 59 percent in the 1990-style forms. This difference may result from dropping the word "origin" from the question. A subsequent experiment has confirmed that question wording differences account for most of the difference in detailed Hispanic reporting, although the examples also contributed (Martin et. al., 2003).

Race Reporting. Race reporting was also affected by differences between the questionnaires. Overall, reports of two or more races more than doubled (.82 percentage points to 2.03 percentage points) in response to the "mark one or more" instruction, there were significantly more reports of Native Hawaiian and Other Pacific Islander (NHOPI), and significantly fewer reports of "Some other race." (These results are not shown.) Contrary to what might have been expected, there is little evidence that allowing respondents to report more than one race reduced single race reporting in the five major race categories. The effects of the questionnaire differences on race reporting by Hispanics were marked. Reporting as White was higher by about 10 percent, while reporting as Some other race was lower by about the same amount, in 2000-style forms, as shown in Table 7. Missing or uncodable responses are excluded,

so the distribution approximates the distribution that would be obtained were missing data imputed. The form differences in reporting of Some other race are consistent with prior research, and are probably due to the effects of reversing the order of Hispanic and race items, as well as the new "one or more" option.

Example Effects: Race Reporting³

In the 1990-style form, examples of "other Asian or Pacific Islander" groups were placed in the leftmost column of the matrix, below the race question (see Fig. 5). These examples (Hmong, Fijian, Laotian, Thai, Tongan, Pakistani, Cambodian) were dropped in the 2000-style form. Table 8 shows the fraction who reported in a race example group in 1990- and 2000-style questionnaires. (People for whom the race question was left entirely blank are dropped from the table.)

The table shows a highly significant form difference (t=3.58) but its direction is unexpected: there are three times as many reports of the example groups in the 2000-style form, which did not list examples. One would expect the use of examples to be associated with higher, not lower, reporting of example categories.

Table 9 shows the fraction reporting in each specific "example group." In general, the 2000-style form elicited more reports of both the Asian and the Pacific Islander example groups, although only the overall differences for Asians and for Pacific Islanders are statistically significant at the .05 level. Note that there were no write-ins of the example Pacific Islander groups in 1990-style forms. Clearly, for the purpose of assessing example

³ This section was not included in Martin (2002).

Table 9.
Percent Writing in Each "Example Group" by Form Type

	1990-style	2000-style	t ₂₀₀₀₋₁₉₉₀
Asian examples106 (.0351)	.300 (.0545)	*2.99
Cambodian014 (.0139)	.041 (.0164)	1.27
Hmong002 (.0023)	.044 (.0243)	†1.69
Pakistani029 (.0217)	.095 (.0317)	†1.72
Thai051 (.0218)	.054 (.0188)	.09
Lao010 (.0097)	.066 (.0276)	†1.92
Pacific Islander examples	0	.057 (.0264)	*2.16
Fijian	0	.019 (.0119)	1.60
Tongan	0	.038 (.0236)	1.61
Another race written in or checked	99.894 (.0351)	99.644 (.0606)	*-3.58
Total percent	100.000	100.000	

*p<.05,2-tailed test.

†p<.10. 2-tailed test

effects for Pacific Islanders, a larger sample is needed. Nevertheless, the difference is consistent for all the groups, and marginally significant for several (t=1.645 is significant at p<.10 with a 2-tailed test), despite very small cell sizes.

It is useful to interpret these results in the context of overall questionnaire differences in reporting of major race groups (from Table 5, Martin, 2002):

- There was no difference between questionnaires in the overall fraction reporting an Asian race: 4.04 percent reported as Asian in 2000-style questionnaires, compared to 4.06 percent in 1990-style questionnaires.
- There was a significant questionnaire difference in the overall fraction reporting as Native Hawaiian and Other Pacific Islander: 0.17 percent reported as NHOPI in 2000-style forms, compared to 0.05 percent in 1990-style forms. Thus, the dif-

ference in reporting of example groups is consistent with an overall reporting difference for this group.

Other questionnaire features, such as splitting the Asian and Pacific Islander (API) category into two separate categories, undoubtedly influenced the results for Pacific Islanders. The Pacific Islander category is probably more populated in 2000-style forms because it is easier for Pacific Islanders to report with the Pacific Islander boxes grouped together rather than interspersed among Asian boxes, as they are in the 1990-style form, and with their own "Other Pacific Islander" check box associated with a write-in space (Cf. Figs. 4 and 5). For these reasons, and because of the very small sample size, we cannot draw even tentative inferences about the effects of the Pacific Islander examples.

The evidence appears stronger that the Asian examples may have

affected reporting. The greater reporting of Asian example groups in the 2000-style form is not consistent with an overall increase in reporting for Asians as a whole, as is the case for Pacific Islanders. Moreover, the form differences are consistent and statistically significant overall, as well as for three individual example groups (Hmong, Pakistani, Lao). The results suggest, although they do not prove, that the use of Asian examples in the 1990-style questionnaire somehow reduced reporting in the example groups.

The contrast between the effects of examples in the Hispanic origin and race items merits further analysis and research. The examples in the 1990 Hispanic origin question seem to have helped to clarify the intent of the question to collect detailed Hispanic origin, while the examples in the race question did not help reporting and may have adversely affected reporting of Asian example groups. Possibly, the different placement of the examples was a factor. In the 1990-style form, the Hispanic examples were prominently placed above the write-in space, just below the "other Spanish/Hispanic" response option. The race examples were off to the left, below the question and remote from the write-in space. Respondents may not have understood how to interpret the meaning of the arrow that connected them. The physical distance between the examples and the write-in space may have meant that many respondents never saw the examples, while those who did may not have realized they were meant to be associated with a write-in space. The examples may have distracted respondents and disrupted the response process. Perhaps respondents were confused by the label "Other API" for

the write-in box in the 1990-style form, which they may not have realized applied to them, so attempted to find some way of reporting their race other than writing in a response. Possibly, respondents in the example groups who found the examples circled them, without writing in a response. Testing these hypotheses and speculations would require examining the forms, which we have not undertaken, or additional experiments. Alternatively, it is possible that features of the questionnaire other than examples somehow explain these results.

In any case, these results suggest the possibility that in some circumstances, examples may interfere with responding in the example

groups. More controlled experiments are needed to understand the effects of examples on race and Hispanic reporting, and how questionnaire features such as their location influence the response process. Meanwhile, considerable caution in their use appears to be warranted.

Conclusions. Additional research is needed into methodological influences on race reporting by Hispanics and non-Hispanics (including experimental research to evaluate mode differences in reporting, which are not explored here but are troubling; see Martin and Gerber, 2003) .

The results raise doubts about the meaningfulness of race data, espe-

cially for Hispanics, for whom race reporting is highly vulnerable to methodological influences. Research is needed to develop more robust race measurement methods that are less vulnerable to methodological effects, especially for Hispanics, and to evaluate the effects of examples on race and Hispanic origin reporting.

Coding, pre-editing, editing, and imputation procedures may substantially affect the quality and comparability of race and Hispanic origin data, and their effects are largely hidden. They need to be documented and evaluated in conjunction with questionnaire design changes.

This page intentionally left blank.

4. Implications and Recommendations

The results of the three experiments share certain themes in common. In this section, we discuss these themes and recommend new avenues for research.

The central lesson of this research is that we must attend to the complex relationships between format and meaning in self-administered questionnaires. These factors have measurable effects on the data. Our experiments demonstrate the effects of questionnaire format on performance on branching instructions, item nonresponse on the household count question and coverage of Hispanics, race and ethnicity reporting, and response rates. In addition, the same questionnaire design feature (e.g., examples) may have complex and varying effects, depending on how it is used and integrated into the questionnaire as a whole. Examples apparently helped respondents understand the intent of the Hispanic origin question, and provide more detailed data. But Asian examples in the race question may have actually reduced reporting of the example groups. **Our results provide a caution that seemingly minor questionnaire variations in the design of a questionnaire can have substantial effects on the data.**

We recommend against introducing any questionnaire changes without fully pretesting and field testing them, since they may have complex and unexpected consequences.

We also note that cognitive testing alone is not sufficient to understand the effects of questionnaire design changes on the data. Cognitive testing alone may prove misleading, and must be supplemented by field experiments. As discussed above, some respondents in cognitive tests, as well as expert reviewers, found the arrows in the Detection method in Experiment A distracting, yet this method resulted in significantly lower error rates.

These results should focus our attention on the importance of the visual aspects of self-administered questionnaires. As the Redline experiment points out, these visual aspects impart meaning in themselves, in that they guide basic questionnaire behavior of respondents. Two factors appear to be at work. One is the attention of respondents. There is little value in altering the wording of a question or instructions if the visual format does not draw the respondent's attention. The second is pattern recognition, or respondents' perception that certain visual elements of a questionnaire are associated, and the inferences and meaning they draw from the groupings they perceive. By manipulating visual features of the questionnaire, Redline succeeds in reducing the frequency with which respondents answer questions they should have skipped, or skip questions they should have answered. **We recommend that the Census Bureau use the Detection method developed by Redline in mail questionnaires because**

it results in fewer errors. The Martin experiment also suggests that visual elements of the design influence respondents' behavior in complex ways, since the effects of examples appear to depend on their placement on the questionnaire (and possibly on other, as yet unidentified, questionnaire design features as well). Gerber attempts a holistic design that graphically integrates the list of residence instructions and the household count box. While we cannot be sure whether the graphical features or the wording changes in her design were responsible for the improvements she obtained, her experiment also reduces errors, supporting its use as well as further research and development of her innovations. **We recommend that the Census Bureau conduct further tests of graphical design and wording changes in the residence instructions, which show promise of improving coverage.**

All three experiments suggest that visual presentation and organization affect respondents' perception, comprehension and performance on self-administered questionnaires. **We recommend further theoretical and empirical questionnaire design research that draws on relevant scientific disciplines to better understand how graphical design features affect the response process.** This includes additional research on the reading of questionnaires. New methodologies, such as eye movement research, are necessary to establish a base-

line for how questionnaires are normally read, in order to be able to better understand how questionnaire design features affect the reading and response process. **We also recommend that the Census Bureau add the analysis of errors of commission and omission to its arsenal of data quality measures, as these are more direct measures of navigational performance than item non-response rates or edit changes.** Using measures of data quality that are relevant both to questionnaire designers and to statisticians may improve their communication about data quality.

Some format effects appear to be more relevant for one group than others. Gerber found that the revised residence rules had stronger effects for Hispanics than non-Hispanics in low coverage areas, perhaps because the changes in wording and organization made the question easier to understand and follow. Martin found response rate differences for the low coverage stratum that included more Black and Hispanic households. In both cases, design changes intended to be "user-friendly" had a bigger effect for respondents who otherwise were less likely to respond or have difficulty with the questionnaire. In addition, questionnaire design differences sometimes affected race reporting by Hispanics and non-Hispanics in opposite ways (see Martin, 2002).

We have not explored the basis for these differences, so cannot address their causes. Possibly, differences in learning and expectations, and training in the form-filling task, underlie them. This suggests that **additional basic research should be carried out to understand how cognitive**

as well as sociocultural factors affect questionnaire performance.

The Martin experiment identifies questionnaire effects that confound comparisons of 1990 and 2000 census data. The degree of confounding cannot be inferred directly from the experimental evidence, which is restricted to mail short forms and does not employ fully edited data. However, we can infer that the differences in the design of 1990 and 2000 mail short forms would have resulted in an increase from the 1990 to the 2000 census in Hispanics' reporting of White race, and a decline in their reporting of detailed Hispanic groups, in *the absence of true change in the racial or ethnic composition or identifications of the population*. The percentage of Hispanics who reported as White (alone) was 51.7 in the 1990 census and 47.9 in Census 2000 (U. S. Census Bureau, 2001). Because the questionnaire effect would have led more Hispanics to report as White, we infer that the decline in White reporting would have been even larger had the 2000-style questionnaire not increased Hispanics' reporting as White, compared to a 1990-style questionnaire. We can also infer that any measured decline from the 1990 to the 2000 census in reporting of detailed Hispanic origins is overstated; the decline would be less if the 2000-style questionnaire had not resulted in less detailed reporting. **The confounding effects of questionnaire differences should be taken into account when comparing 1990 and 2000 census data.**

This panel of the AQE had as its main purpose calibrating the measurement properties of the 1990 and 2000 mail short form questionnaires. By replicating the 1990

census form in Census 2000, it is possible to evaluate whether differences in reporting are attributable to the questionnaire, and must be taken into account in interpreting population trends from one census to the next. The results demonstrate that replicating a prior census's questions in the current census can help shed light on possible errors and reporting differences that otherwise would be the subject of unchecked conjecture and speculation. For example, the AQE data made it possible to understand how changes in the questionnaire caused an unexpected decline in detailed Hispanic reporting between 1990 and 2000 censuses. Without the AQE data, analysts might have been tempted to explain the apparent trend as the result of an increasing pan-Hispanic identification in the Hispanic population, when it appears to be an artifact of the questionnaire design. **We recommend that a similar replication be carried out in 2010, and that the long form be included,** to calibrate the different versions of the long form used in the census in 2000 and in the American Community Survey (ACS) in 2010.

The AQE sample was not large enough to permit us to test design changes separately. Rather, the experiments (particularly the Martin and Gerber experiments) all tested "packages" which confound the effects of format and content, and which do not permit us to isolate the factors responsible for effects on particular populations. More sophisticated experimental designs might permit future experiments to disentangle these effects. In addition, where coverage is measured, larger reinterview samples are needed, since omissions and erroneous enumerations account for a very small fraction of

the entire population. These limitations were implicit in the size and scope of the current AQE. Both for the purposes of calibrating measurements (as in the Martin

experiment), and the purpose of measuring small improvements with high precision (as in the Gerber experiment), **we recommend much larger sample**

sizes and more sophisticated experimental designs for Alternate Questionnaire Experiments of the future.

This page intentionally left blank.

References

- Bates, N. A. (1991), "The Alternative Questionnaire Experiment: Preliminary Report of the 100-percent Items." 1990 Decennial Census Preliminary Research and Evaluation Memorandum No. 108.
- Bates, N.A., (1992), "Additional Results from the 1990 Alternative Questionnaire Experiment." 1990 Decennial Census Preliminary Research and Evaluation Memorandum No. 146.
- Dajani, A. and Scaggs, M. A. (2001) "AQE 2000 Response Rate Analysis." Unpublished memorandum, Census Bureau.
- Fay, R. E. (1998) VPLX Program Documentation, Vol. 1. Census Bureau.
- Gerber, E., Dajani, A., and Scaggs, M. A. (2002) An Experiment to Improve Coverage Through Revised Roster Instructions. Final report of Census 2000 Alternative Questionnaire Experiment. U.S. Census Bureau.
- Martin, E. (2002) Questionnaire Effects on Reporting of Race and Hispanic Origin: Results of a Replication of the 1990 Mail Short Form in Census 2000. Final report of Census 2000 Alternative Questionnaire Experiment. U.S. Census Bureau.
- Martin, E., Sheppard, D., Bentley, M., and Bennett, C. 2003. "Results of 2003 National Census Test of Race and Hispanic Questions." Report presented to Census Bureau Race and Ethnic Advisory Committees, October 1, 2003.
- Martin, E. and Gerber, E. (2003) "Methodological Influences on Comparability of Race Measurements: Several Cautionary Examples." Paper presented at the American Statistical Association, San Francisco, CA.
- McKenney, N. D., Fernandez, E. W. and Masamura, W. T. (1985) "The quality of the race and Hispanic origin information reported in the 1980 census." Proceedings of the American Statistical Association (Survey Research Methods Section), pp. 46-50.
- Mockovak, W. (1984) Alternative Questionnaires Experiment. Preliminary Evaluations Results Memorandum No. 93, 1980 Census. U.S. Census Bureau.
- Redline, C., Dillman, D., Dajani, A. and Scaggs, M. A. (2002) Effects of Altering The Design of Branching Instructions On Navigational Performance in Census 2000. Final report of Census 2000 Alternative Questionnaire Experiment. U.S. Census Bureau.
- Singer, P. and Ennis, S. R. (2003) Content Reinterview Survey of the Census 2000: Accuracy of Data for Selected Population and Housing Characteristics as Measured by Reinterview. Census 2000 Evaluation B5. U. S. Census Bureau.
- U. S. Census Bureau (2001) "Population by Race and Hispanic or Latino Origin for the United States: 1990 and 2000." Census 2000 PHC-T-1.

This page intentionally left blank.

Appendix

Figure 1. Illustration of the five branching instruction treatments

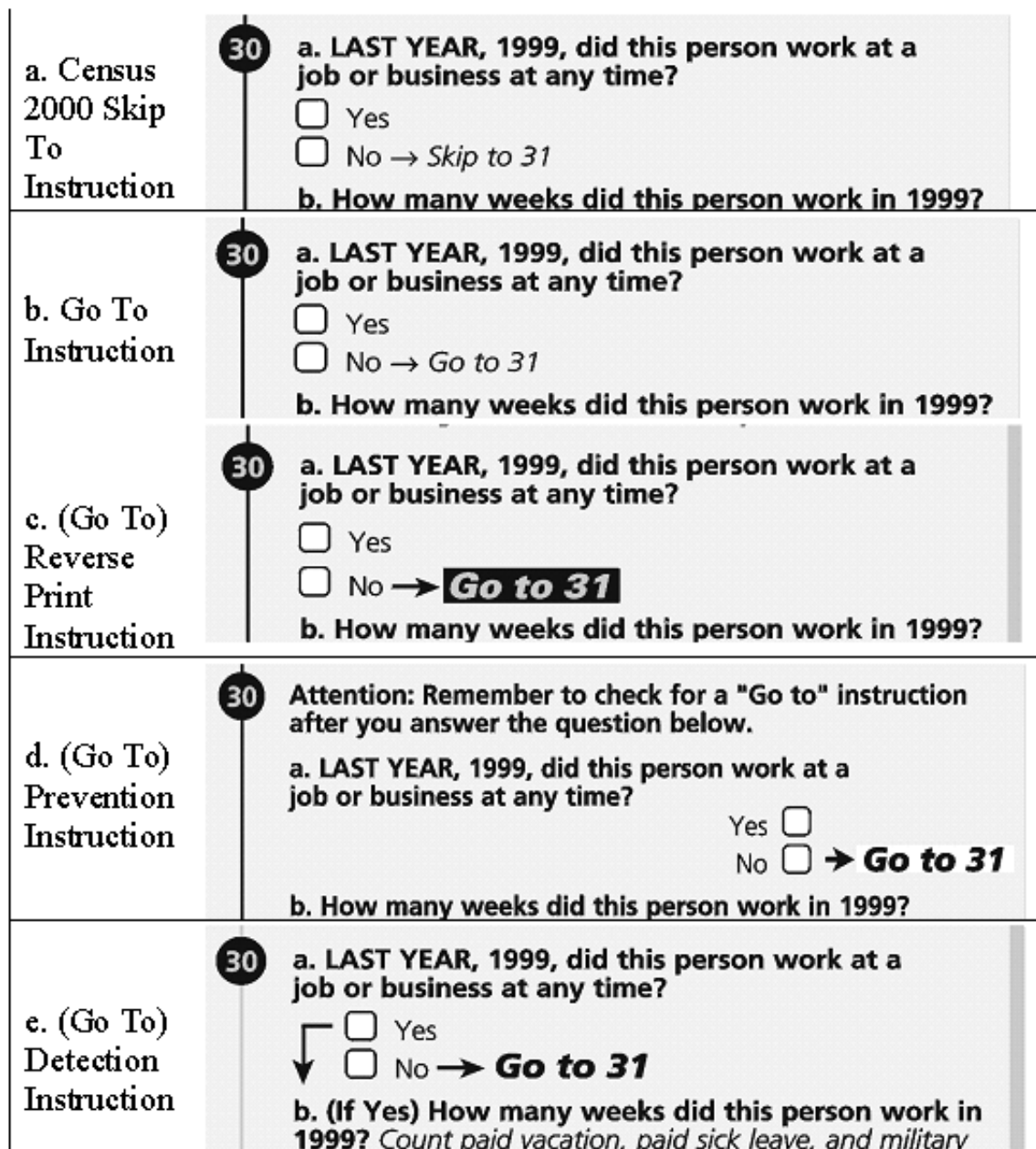


Figure 2. Census 2000 Roster Instructions (Experiment B)

United States
Census
2000

U.S. Department of Commerce • Bureau of the Census

This is the official form for all the people at this address. It is quick and easy, and your answers are protected by law. Complete the Census and help your community get what it needs — today and in the future!

Start Here

Please use a black or blue pen.

1. How many people were living or staying in this house, apartment, or mobile home on April 1, 2000?

Number of people

INCLUDE in this number:

- foster children, roomers, or housemates
- people staying here on April 1, 2000 who have no other permanent place to stay
- people living here most of the time while working, even if they have another place to live

DO NOT INCLUDE in this number:

- college students living away while attending college
- people in a correctional facility, nursing home, or mental hospital on April 1, 2000
- Armed Forces personnel living somewhere else
- people who live or stay at another place most of the time

2. Is this house, apartment, or mobile home — Mark **ONE** box.

Owned by you or someone in this household with a mortgage or loan?

Owned by you or someone in this household free and clear (without a mortgage or loan)?

Rented for cash rent?

Occupied without payment of cash rent?

3. Please answer the following questions for each person living in this house, apartment, or mobile home. Start with the name of one of the people living here who owns, is buying, or rents this house, apartment, or mobile home. If there is no such person, start with any adult living or staying here. We will refer to this person as Person 1.

What is this person's name? *Print name below.*

Last Name MI

First Name

4. What is Person 1's telephone number? *We may call this person if we don't understand an answer.*

Area Code + Number

5. What is Person 1's sex? Mark **ONE** box.

Male Female

6. What is Person 1's age and what is Person 1's date of birth?

Age on April 1, 2000

Print numbers in boxes.

Month Day Year of birth

→ **NOTE: Please answer BOTH Questions 7 and 8.**

7. Is Person 1 Spanish/Hispanic/Latino? Mark the "No" box if **not** Spanish/Hispanic/Latino.

No, not Spanish/Hispanic/Latino Yes, Puerto Rican

Yes, Mexican, Mexican Am., Chicano Yes, Cuban

Yes, other Spanish/Hispanic/Latino — *Print group.* ↗

8. What is Person 1's race? Mark **one or more races** to indicate what this person considers himself/herself to be.

White

Black, African Am., or Negro

American Indian or Alaska Native — *Print name of enrolled or principal tribe.* ↗

Asian Indian Japanese Native Hawaiian

Chinese Korean Guamanian or Chamorro

Filipino Vietnamese Samoan

Other Asian — *Print race.* ↗ Other Pacific Islander — *Print race.* ↗

Some other race — *Print race.* ↗

→ **If more people live here, continue with Person 2.**

OMB No. 0607-0862: Approval Expires 07/31/2002

Form **S-800A.1**

Figure 3. Roster Instructions in Experimental Form (Experiment B)

United States Census 2000

U.S. Department of Commerce • Bureau of the Census

This is the official form for all the people at this address. It is quick and easy, and your answers are protected by law. Complete the Census and help your community get what it needs — today and in the future!

Before you answer Question 1, count the people living at your address using our guidelines:

Include these people:

- People who live here most of the time, even if they have somewhere else to live
- Roommates or boarders
- Children living here, including foster children
- People staying here on April 1, 2000 who have no other permanent place to stay

Do NOT include these people:

- College students who live away
- Armed Forces personnel who live away
- People who live or stay somewhere else most of the time
- People who, on April 1, 2000, are in:
 - a nursing home
 - a mental hospital
 - a correctional facility (example: jail)

1. How many people were living or staying in this house, apartment, or mobile home on April 1, 2000?

Number of people =

2. Is this house, apartment, or mobile home —
Mark **ONE** box.

Owned by you or someone in this household with a mortgage or loan?

Owned by you or someone in this household free and clear (without a mortgage or loan)?

Rented for cash rent?

Occupied without payment of cash rent?

3. Please answer the following questions for each of the persons counted in Question 1. Start with the name of one of the people living here who owns, is buying, or rents this place. If there is no such person, start with any adult living here. We will refer to this person as Person 1.

What is this person's name? Print name below.

Last Name

First Name MI

4. What is Person 1's telephone number? We may call this person if we don't understand an answer.

Area Code + Number

- -

5. What is Person 1's sex? Mark **ONE** box.

Male Female

6. What is Person 1's age and what is Person 1's date of birth?

Age on April 1, 2000

Print numbers in boxes.

Month Day Year of birth

→ **NOTE: Please answer BOTH Questions 7 and 8.**

7. Is Person 1 Spanish/Hispanic/Latino? Mark the "No" box if **not** Spanish/Hispanic/Latino.

No, not Spanish/Hispanic/Latino Yes, Puerto Rican

Yes, Mexican, Mexican Am., Chicano Yes, Cuban

Yes, other Spanish/Hispanic/Latino — *Print group.* ↗

8. What is Person 1's race? Mark **one or more races** to indicate what this person considers himself/herself to be.

White

Black, African Am., or Negro

American Indian or Alaska Native — *Print name of enrolled or principal tribe.* ↗

Asian Indian Japanese Native Hawaiian

Chinese Korean Guamanian or Chamorro

Filipino Vietnamese Samoan

Other Asian — *Print race.* ↗ Other Pacific Islander — *Print race.* ↗

Some other race — *Print race.* ↗

→ **If more people live here, continue with Person 2.**

OMB No. 0607-0862: Approval Expires 07/31/2002

Form **S-800A.3**

Figure 4. 2000-style form: Race and Hispanic Questions (Experiment C)

Your answers are important!
Every person in the Census counts.

Person 2

1. What is Person 2's name? *Print name below.*

Last Name

First Name

MI

2. How is this person related to Person 1? *Mark ONE box.*

<input type="checkbox"/> Husband/wife	If NOT RFI ATFD to Person 1:
<input type="checkbox"/> Natural-born son/daughter	<input type="checkbox"/> Roomer, boarder
<input type="checkbox"/> Adopted son/daughter	<input type="checkbox"/> Housemate, roommate
<input type="checkbox"/> Stepson/stepdaughter	<input type="checkbox"/> Unmarried partner
<input type="checkbox"/> Brother/sister	<input type="checkbox"/> Foster child
<input type="checkbox"/> Father/mother	<input type="checkbox"/> Other nonrelative
<input type="checkbox"/> Grandchild	
<input type="checkbox"/> Parent in law	
<input type="checkbox"/> Son in law/daughter in law	
<input type="checkbox"/> Other relative. <i>Print exact relationship.</i> →	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>

3. What is this person's sex? *Mark ONE box.*

Male Female

4. What is this person's age and what is this person's date of birth? *Print numbers in boxes.*

Age on April 1, 2000 Month Day Year of birth

→ **NOTE: Please answer BOTH Questions 5 and 6.** →

5. Is this person Spanish/Hispanic/Latino? *Mark the "No" box if not Spanish/Hispanic/Latino.*

<input type="checkbox"/> No, not Spanish/Hispanic/Latino	<input type="checkbox"/> Yes, Puerto Rican
<input type="checkbox"/> Yes, Mexican, Mexican Am., Chicano	<input type="checkbox"/> Yes, Cuban
<input type="checkbox"/> Yes, other Spanish/Hispanic/Latino — <i>Print group.</i> ↘	

6. What is this person's race? *Mark one or more races to indicate what this person considers himself/herself to be.*

White

Black, African Am., or Negro

American Indian or Alaska Native: *Print name of enrolled or principal tribe.* ↘

<input type="checkbox"/> Asian Indian	<input type="checkbox"/> Japanese	<input type="checkbox"/> Native Hawaiian
<input type="checkbox"/> Chinese	<input type="checkbox"/> Korean	<input type="checkbox"/> Guamanian or Chamorro
<input type="checkbox"/> Filipino	<input type="checkbox"/> Vietnamese	<input type="checkbox"/> Samoan
<input type="checkbox"/> Other Asian — <i>Print race.</i> ↘	<input type="checkbox"/> Other Pacific Islander — <i>Print race.</i> ↘	

Some other race — *Print race.* ↘

→ **If more people live here, continue with Person 3.** →

Figure 5. 1990-style form: Race and Hispanic Questions (Experiment C)

Please fill one column only for each person listed in Question 1.	PERSON 1		PERSON 2		
	Last name		Last name		Last name
	First name	Middle initial	First name	Middle initial	First name
<p>2. How is this person related to PERSON 1?</p> <p>Fill ONE circle for each person.</p> <p>If Other relative of person in column 1, fill circle and print exact relationship, such as mother-in-law, grandparent, son-in-law, niece, cousin, and so on.</p>	<p>START in this column with the household member (or one of the members) in whose name the home is owned, being bought, or rented.</p> <p>If there is no such person, start in this column with any adult household member.</p>		<p>If a RELATIVE of Person 1:</p> <p><input type="checkbox"/> Husband/wife <input type="checkbox"/> Brother/sister</p> <p><input type="checkbox"/> Natural-born or adopted son/daughter <input type="checkbox"/> Father/mother</p> <p><input type="checkbox"/> Grandchild <input type="checkbox"/> Other relative</p> <p><input type="checkbox"/> Stepson/stepdaughter</p> <p>If NOT RELATED to Person 1:</p> <p><input type="checkbox"/> Roomer, boarder, or foster child <input type="checkbox"/> Unmarried partner</p> <p><input type="checkbox"/> Housemate, roommate <input type="checkbox"/> Other nonrelative</p>		<p>If a RELATIVE of Person 1:</p> <p><input type="checkbox"/> Husband/wife</p> <p><input type="checkbox"/> Natural-born or adopted son/daughter</p> <p><input type="checkbox"/> Stepson/stepdaughter</p> <p>If NOT RELATED to Person 1:</p> <p><input type="checkbox"/> Roomer, boarder, or foster child</p> <p><input type="checkbox"/> Housemate, roommate</p>
<p>3. Sex</p> <p>Fill ONE circle for each person.</p>	<p><input type="checkbox"/> Male <input type="checkbox"/> Female</p>	<p><input type="checkbox"/> Male <input type="checkbox"/> Female</p>	<p><input type="checkbox"/> Male <input type="checkbox"/> Female</p>	<p><input type="checkbox"/> Male</p>	
<p>4. Race</p> <p>Fill ONE circle for the race that the person considers himself/herself to be.</p> <p>If Indian (Amer.), print the name of the enrolled or principal tribe.</p> <p>If Other Asian or Pacific Islander (API), print one group, for example: Hmong, Fijian, Laotian, Thai, Tongan, Pakistani, Cambodian, and so on.</p> <p>If Other race, print race.</p>	<p><input type="checkbox"/> White</p> <p><input type="checkbox"/> Black or Negro</p> <p><input type="checkbox"/> Indian (Amer.) (Print the name of the enrolled or principal tribe.)</p> <p><input type="checkbox"/> Eskimo</p> <p><input type="checkbox"/> Aleut</p> <p><input type="checkbox"/> Asian or Pacific Islander (API)</p> <p><input type="checkbox"/> Chinese <input type="checkbox"/> Japanese</p> <p><input type="checkbox"/> Filipino <input type="checkbox"/> Asian Indian</p> <p><input type="checkbox"/> Hawaiian <input type="checkbox"/> Samoan</p> <p><input type="checkbox"/> Korean <input type="checkbox"/> Guamanian</p> <p><input type="checkbox"/> Vietnamese <input type="checkbox"/> Other API</p> <p><input type="checkbox"/> Other race (Print race)</p>	<p><input type="checkbox"/> White</p> <p><input type="checkbox"/> Black or Negro</p> <p><input type="checkbox"/> Indian (Amer.) (Print the name of the enrolled or principal tribe.)</p> <p><input type="checkbox"/> Eskimo</p> <p><input type="checkbox"/> Aleut</p> <p><input type="checkbox"/> Asian or Pacific Islander (API)</p> <p><input type="checkbox"/> Chinese <input type="checkbox"/> Japanese</p> <p><input type="checkbox"/> Filipino <input type="checkbox"/> Asian Indian</p> <p><input type="checkbox"/> Hawaiian <input type="checkbox"/> Samoan</p> <p><input type="checkbox"/> Korean <input type="checkbox"/> Guamanian</p> <p><input type="checkbox"/> Vietnamese <input type="checkbox"/> Other API</p> <p><input type="checkbox"/> Other race (Print race)</p>	<p><input type="checkbox"/> White</p> <p><input type="checkbox"/> Black or Negro</p> <p><input type="checkbox"/> Indian (Amer.)</p> <p><input type="checkbox"/> Eskimo</p> <p><input type="checkbox"/> Aleut</p> <p><input type="checkbox"/> Asian</p> <p><input type="checkbox"/> Chinese</p> <p><input type="checkbox"/> Filipino</p> <p><input type="checkbox"/> Hawaiian</p> <p><input type="checkbox"/> Korean</p> <p><input type="checkbox"/> Vietnamese</p> <p><input type="checkbox"/> Other</p>		
<p>5. Age and year of birth</p> <p>a. Print each person's age at last birthday. Fill in the matching circle below each box.</p> <p>b. Print each person's year of birth and fill the matching circle below each box.</p>	<p>a. Age</p> <p>0 v 0 v 0 v</p> <p>1 v 1 v 1 v</p> <p>2 v 2 v</p> <p>3 v 3 v</p> <p>4 v 4 v</p> <p>5 v 5 v</p> <p>6 v 6 v</p> <p>7 v 7 v</p> <p>8 v 8 v</p> <p>9 v 9 v</p>	<p>b. Year of birth</p> <p>1 v 0 v 0 v 0 v</p> <p>2 v 8 v 1 v 1 v</p> <p>9 v 2 v 2 v</p> <p>3 v 3 v</p> <p>4 v 4 v</p> <p>5 v 5 v</p> <p>6 v 6 v</p> <p>7 v 7 v</p> <p>8 v 8 v</p> <p>9 v 9 v</p>	<p>a. Age</p> <p>0 v 0 v 0 v</p> <p>1 v 1 v 1 v</p> <p>2 v 2 v</p> <p>3 v 3 v</p> <p>4 v 4 v</p> <p>5 v 5 v</p> <p>6 v 6 v</p> <p>7 v 7 v</p> <p>8 v 8 v</p> <p>9 v 9 v</p>	<p>b. Year of birth</p> <p>1 v 0 v 0 v 0 v</p> <p>2 v 8 v 1 v 1 v</p> <p>9 v 2 v 2 v</p> <p>3 v 3 v</p> <p>4 v 4 v</p> <p>5 v 5 v</p> <p>6 v 6 v</p> <p>7 v 7 v</p> <p>8 v 8 v</p> <p>9 v 9 v</p>	<p>a. Age</p> <p>0 v 0</p> <p>1 v 1</p> <p>2</p> <p>3</p> <p>4</p> <p>5</p> <p>6</p> <p>7</p> <p>8</p> <p>9</p>
<p>6. Is this person of Spanish/Hispanic origin?</p> <p>Fill ONE circle for each person.</p> <p>If Yes, other Spanish/Hispanic, print one group.</p>	<p><input type="checkbox"/> No (not Spanish/Hispanic)</p> <p><input type="checkbox"/> Yes, Mexican, Mexican-Am., Chicano</p> <p><input type="checkbox"/> Yes, Puerto Rican</p> <p><input type="checkbox"/> Yes, Cuban</p> <p><input type="checkbox"/> Yes, other Spanish/Hispanic (Print one group, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)</p>	<p><input type="checkbox"/> No (not Spanish/Hispanic)</p> <p><input type="checkbox"/> Yes, Mexican, Mexican-Am., Chicano</p> <p><input type="checkbox"/> Yes, Puerto Rican</p> <p><input type="checkbox"/> Yes, Cuban</p> <p><input type="checkbox"/> Yes, other Spanish/Hispanic (Print one group, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)</p>	<p><input type="checkbox"/> No (not Spanish/Hispanic)</p> <p><input type="checkbox"/> Yes, Mexican, Mexican-Am., Chicano</p> <p><input type="checkbox"/> Yes, Puerto Rican</p> <p><input type="checkbox"/> Yes, Cuban</p> <p><input type="checkbox"/> Yes, other Spanish/Hispanic (Print one group, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.)</p>		

Form S-800A.2

