

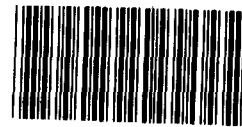
GAO

Report to the Chairman, Subcommittee
on Census, Statistics, and Postal
Personnel, Committee on Post Office
and Civil Service, House of
Representatives

January 1993

CENSUS REFORM

Early Outreach and Decisions Needed on Race and Ethnic Questions



148582

**RESTRICTED--Not to be released outside the
General Accounting Office unless specifically
approved by the Office of Congressional
Relations.**

556 331 RELEASED

General Government Division

B-249774

January 28, 1993

The Honorable Thomas C. Sawyer
Chairman, Subcommittee on Census,
Statistics, and Postal Personnel
Committee on Post Office and Civil Service
House of Representatives

Dear Mr. Chairman:

This report responds to your request for information on (1) the development of the race and ethnic questions on the 1990 decennial census, (2) the quality of the race and ethnic data from the decennial census, and (3) the status of the Bureau of the Census' research programs for these questions for the 2000 census. This information should help the Subcommittee as it oversees the Bureau's progress in preparing for the 2000 census.¹ On December 15, 1992, we briefed the Subcommittee on our findings. This report documents and supplements the information we provided at the briefing.

Results in Brief

Race and ethnic questions are among the most technically complex and publicly controversial questions asked on the decennial census. The experiences from the 1990 census showed that the Bureau needs to begin early in the decade to work with a diverse group of customers—including Congress, other federal agencies, researchers, and organizations representing the interests of various race and ethnic groups—to identify the data needs for the 2000 census and the best ways these needs can be met. For the 1990 census, the Bureau was not able to build the necessary consensus for its recommended version of the race question in spite of a special research and testing program on the race and ethnic questions. As a result, the final format of the race question was decided late in the decade after protracted debate and was contrary to the Bureau's initial recommendations. A discussion of the evolution of the race and ethnic questions appears in appendix II.

Bureau evaluations suggest that the data from the 1990 race and Hispanic origin questions² are generally of high quality. However, these evaluations also suggest that several data quality problems that confronted the Bureau

¹See *Decennial Census: 1990 Results Show Need for Fundamental Reform* (GAO/GGD-92-94, June 9, 1992), which provides an overview of the major lessons learned from the 1990 census and identifies opportunities for fundamental reform.

²Although the term, "Spanish/Hispanic origin" appears on the census questionnaire, this report uses the term "Hispanic origin" for simplicity.

after the 1980 census continue to plague the Bureau. For example, Hispanics, particularly foreign-born Hispanics, have difficulty classifying themselves by race. Non-Hispanics, on the other hand, tend to skip over the Hispanic origin question when they should indicate that they are not of Hispanic origin. These two phenomena create data quality problems for both the race and Hispanic origin questions. In addition, the number of persons reported as "Other race," while only about 4 percent of the total U.S. population, continues to grow at a much faster rate than the total population. Data quality issues are discussed in greater detail in appendix III.

The Bureau faces a difficult balancing act between efforts to simplify the questionnaire and requests from minority populations for identification on the census questionnaire. In addition, federal decisionmakers are pressuring the Bureau to develop ever more current and detailed intercensal race and ethnic data, which are beyond the scope of the decennial census.

The Bureau has developed a plan for research and development of race and ethnic questions for the 2000 census. The plan suggests an awareness of the major issues on race and ethnic questions needing attention, particularly the importance of early and open consultation with interested parties. The major challenge the Bureau faces is developing a truly cooperative working relationship with key customers, particularly representatives of race and ethnic groups, so that the resources the Bureau commits to this important area of research are directed toward improvements that are broadly understood and accepted. Additional detail regarding the status of the Bureau's research on measuring race and ethnicity for the 2000 census is provided in appendix IV.

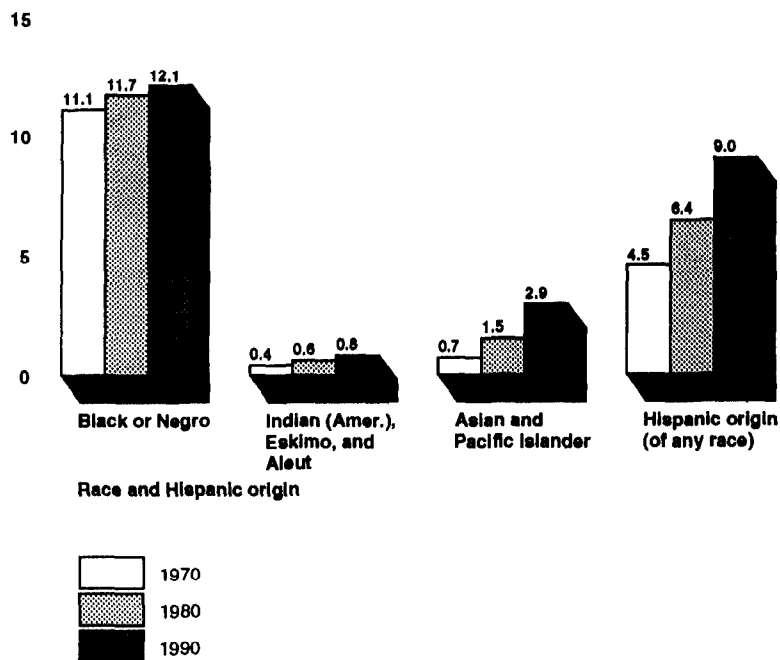
Background

Racial and Ethnic Minority Population Is Growing

As figure 1 shows, the racial and ethnic minority population in the United States has grown. In 1990, American Indians and Alaska Natives, Asian and Pacific Islanders, and Blacks combined made up 15.8 percent of the U.S. population. Hispanic persons, who may be of any race, made up 9 percent of the population.

Figure 1: Growth of Minority Populations

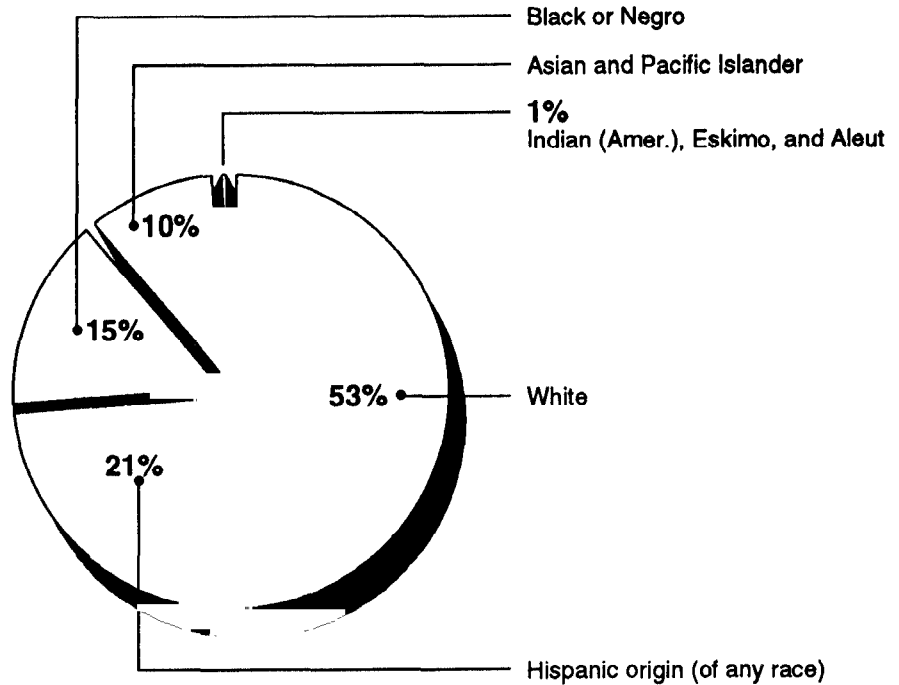
20 Percent of U.S. population



Source: Census Bureau.

Moreover, racial and ethnic minority populations are growing at a faster rate than the White population. The White, not-of-Hispanic-origin population dropped from 83.5 percent of the total population in 1970 to 79.6 percent in 1980 and to 75.6 percent in 1990. This means that close to a quarter of the total U.S. population today is either Hispanic or of a minority race. The Bureau projects that by the middle of the next century, virtually half of the population will be made up of American Indians and Alaska Natives, Asian and Pacific Islanders, Blacks, and Hispanics, as figure 2 illustrates.

Figure 2: Population Projection by Race and Ethnic Group for the Year 2050



To avoid doublecounting, projections by race group do not include persons of Hispanic origin.

Source: Census Bureau.

Tabulating Race and Ethnic Data Is a Multistep Process

The format of the race and Hispanic origin questions as they appeared in the 1990 census is shown in figure 3.

Figure 3: 1990 Census Race and Hispanic Origin Questions

<p>4. Race 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>	<ul style="list-style-type: none"> <input type="radio"/> White <input type="radio"/> Black or Negro <input type="radio"/> Indian (Amer.) (Print the name of the enrolled or principal tribe.) ↗ <div style="border: 1px dashed black; height: 20px; width: 100%; margin-top: 5px;"></div> <input type="radio"/> Eskimo <input type="radio"/> Aleut <li style="text-align: center;"><u>Asian or Pacific Islander (API)</u> <input type="radio"/> Chinese <input type="radio"/> Japanese <input type="radio"/> Filipino <input type="radio"/> Asian Indian <input type="radio"/> Hawaiian <input type="radio"/> Samoan <input type="radio"/> Korean <input type="radio"/> Guamanian <input type="radio"/> Vietnamese <input type="radio"/> Other API ↗ <div style="border: 1px dashed black; height: 20px; width: 100%; margin-top: 5px;"></div> <input type="radio"/> Other race (Print race) ↗ <div style="border: 1px dashed black; height: 20px; width: 100%; margin-top: 5px;"></div>
<p>7. Is this person of Spanish/Hispanic origin? Fill ONE circle for each person.</p> <p>If Yes, other Spanish/Hispanic, print one group. _____ →</p>	<ul style="list-style-type: none"> <input type="radio"/> No (not Spanish/Hispanic) <input type="radio"/> Yes, Mexican, Mexican-Am., Chicano <input type="radio"/> Yes, Puerto Rican <input type="radio"/> Yes, Cuban <input type="radio"/> Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.) ↗ <div style="border: 1px dashed black; height: 20px; width: 100%; margin-top: 5px;"></div>

Source: 1990 Census.

The Bureau undertook several activities to produce race and ethnic data once the decennial census questionnaires were returned. These activities included the following:

- First, categories marked on all questionnaires under the race and ethnic items were read by computer using optical scanning equipment.
- Second, write-in responses for race and Hispanic origin were reviewed for the first time in census history using an automated coding operation in which write-in responses were matched by computer against a master code list. This operation enabled the Bureau to review all write-in responses and thereby improve the quality of the race data. With automated coding, the Bureau was also able to reclassify entries that would otherwise have been misclassified. For example, persons who marked the "Other race" category and wrote "German" were assigned a detailed code by the Bureau so as to capture the write-in information, but they were then reclassified as "White." Almost all write-in entries to the race question were reviewed through this automated process, whereas only a sample of the Hispanic origin questions were reviewed in this fashion.
- Third, the Bureau allocated responses when the person failed to answer the question or failed to answer the question in a complete or legible fashion. Allocation means that the person was assigned the race or Hispanic origin (whichever characteristic was missing) of another household member or neighboring household that was similar in some respect.
- The Bureau also produced a special data file called the "modified race file." This entailed assigning persons in the "Other race" category to a standard race category based on a methodology described later in this report. In this manner, census data were transformed into race categories that corresponded with those used in other federal data sets to meet the needs of federal and state agencies and researchers.

Objectives, Scope, and Methodology

Our objectives were to (1) trace the development of the race and ethnic questions on the 1990 decennial census, (2) identify any quality problems that may exist with race and ethnic data for the 1990 census, and (3) identify the major challenges confronting the Bureau as it begins research on race and ethnic questions for the 2000 census.

To meet the first objective, we examined literature on the changes to the race and ethnic questions between the 1980 and 1990 censuses.³ Two census questions—one on race and the other on Hispanic origin—were our focus because these questions are asked of the entire population. We met with technical staff from the Department of Justice, the Equal

³Our primary source was the Bureau's Content Determination Report: Race and Ethnic Origin (1990 CDR-6, Feb. 1991).

Employment Opportunity Commission, and the Department of Health and Human Services to determine whether census data met their needs and whether the Bureau was generally responsive to their concerns. We chose these agencies because they were identified by the Bureau as major users of census data on race and ethnicity.

To gather perspectives from the major race and ethnic groups, we interviewed members of the census advisory committees for four different populations—American Indian and Alaska Natives, Asian and Pacific Islanders, Blacks, and Hispanics. The advisory committees were established in 1986 by the Department of Commerce to provide a channel of communication between the Bureau and certain minority populations on content, data needs, enumeration, outreach, publicity, and other topics. The members we spoke with included the chairpersons of the respective committees and/or other committee members identified by the chairpersons as knowledgeable about the development of the race and ethnic questions.

We met our second objective by examining Bureau research papers, 1990 census data, and related Bureau evaluations to identify data quality concerns. To meet our third objective, we reviewed Bureau planning documents to gauge progress in preparing for the 2000 census. We interviewed Bureau officials throughout this effort and spoke with experts outside the Bureau. We also reviewed literature on the problems in measuring race and ethnicity.

We did our work from June through November 1992 in Washington, D.C., and at Bureau headquarters in Suitland, MD, in accordance with generally accepted government auditing standards.

Agency Comments

We discussed the issues presented with Bureau staff involved in the Bureau's race and ethnic research program. In general, they agreed with the issues presented and offered suggestions for technical clarifications. We have made these suggested changes where appropriate.

As agreed with your office, unless you publicly release its contents earlier, we plan no further distribution of this report until 30 days from the date of issuance. At that time we will send copies of this report to the Secretary of Commerce; the Director, Bureau of the Census; the Director, Office of

Management and Budget; and interested congressional committees. We will also make copies available to others upon request.

This report was prepared under the direction of J. Christopher Mihm, Assistant Director. Jacquelyn B. Werth was the Evaluator-in-Charge. If you have any questions, please call me on (202) 275-8676.

Sincerely yours,



L. Nye Stevens
Director, Government Business
Operations and Information Issues

Contents

Letter	1
Appendix I Race and Ethnic Data Meet Variety of Pressing Policy and Programmatic Needs	12
Appendix II Evolution of the Race and Ethnic Questions on the 1990 Census	14
Appendix III Quality of Race and Ethnic Data	22
Appendix IV Status of Census Bureau Research on Race and Ethnicity	29
Figures	
Figure 1: Growth of Minority Populations	3
Figure 2: Population Projection by Race and Ethnic Group for the Year 2050	4
Figure 3: 1990 Census Race and Hispanic Origin Questions	5
Figure II.1: 1980 Census Race and Hispanic Origin Questions	15
Figure II.2: Short Version of the Race Question	19
Figure III.1: Response Consistency for Race Question	22
Figure III.2: Response Consistency for Hispanic Origin Question	23
Figure III.3: Comparative Allocation Rates	24
Figure III.4: Growth of Race Groups From 1980 to 1990	26
Figure III.5: Write-In Responses to "Other Race" Category	27

Contents

Abbreviations

EEOC Equal Employment Opportunity Commission
NCHS National Center for Health Statistics

Race and Ethnic Data Meet Variety of Pressing Policy and Programmatic Needs

The race and ethnic data collected by the census serves many important purposes. The enactment of the Civil Rights Act of 1964 (P.L. 88-352) created a need for the federal government to collect statistics on race for compliance and enforcement purposes in such areas as education. Other statutes followed, including the Voting Rights Act of 1965 (P.L. 89-110) and the Equal Employment Opportunity Act of 1972 (P.L. 92-261), the implementation of which depend in part on statistics on race. A desire for data on the status of an emerging Hispanic population created a need for data by ethnic group as well. This resulted in the passage of P.L. 94-311 in 1976, which required certain federal agencies to publish statistics on the social, health, and economic condition of Americans of Hispanic origin.

Statistical Policy Directive No. 15, which has been in effect since 1978, provides standard classifications for race and ethnicity in response to the need for consistent federal statistical data.¹ Enforced by the Office of Management and Budget, this directive defines four racial groups (American Indian or Alaska Native, Asian or Pacific Islander, Black, and White) and one ethnic group (Hispanic), based on geographical or cultural, rather than scientific, origins. Agencies must be able to aggregate all race and ethnic data to these basic categories, although they are also free to collect more detailed information within these categories, which is the practice with the decennial census. The Bureau was granted an exception to Directive No. 15 in order to offer the category of "Other race" for those persons who do not identify with any of the race categories provided. However, the Bureau maintains the capability of reclassifying data into the standard race categories.

Census data on race and ethnicity are used by many federal agencies to fulfill these legislative requirements. The Department of Justice is one of the biggest users of these data. Race and ethnic data are critical, for example, to Justice's efforts to promote fair voting practices by protecting minority group participation. Under the Voting Rights Act, Justice uses census data to determine if the relocation of voting places within a precinct could diminish minority participation. In addition, these data are used by Justice to review state redistricting plans to ensure that the new lines drawn do not weaken the voting influence of any minority group.

¹We recently reviewed federal agencies' definitions of race and ethnicity to determine causes for any inconsistent reporting of data. In Federal Data Collection: Agencies' Use of Consistent Race and Ethnic Definitions (GAO/GGD-93-25, Dec. 15, 1992), we reported that inconsistent use of race and ethnic terms can occur in the reporting of data when agencies use external sources such as state-provided data. Inconsistent use of race and ethnic terms can also occur when people are classified by observer-identification instead of self-identification. The Office of Management and Budget's Directive No. 15 is reproduced in appendix I of that report.

The Equal Employment Opportunity Commission (EEOC) uses census data to advance equal opportunity for members of minority groups. For example, EEOC uses the data to establish and evaluate guidelines for federal affirmative action plans. In addition, EEOC relies upon census data to evaluate employment patterns of minorities in the private sector. More specifically, EEOC uses census data to compare a minority group's representation in an employer's work force with an estimate of group members' availability for these jobs. Census data on employment, industry, and commuting patterns is used to estimate availability. If the representation of minorities in the employer's work force is significantly smaller than the availability estimates, EEOC may use these data to support administrative enforcement and litigation efforts.

Census data on race and ethnicity are also critical to analyses of the health status of the population. The National Center for Health Statistics (NCHS), within the Department of Health and Human Services and the Centers for Disease Control, is required to collect a wide range of health statistics in order to monitor the extent and nature of illness and disability of the population, and the impact of illness and disability on the well-being of the population. NCHS needs population counts by race and ethnic group, as collected in the census, to produce health statistics on these special populations. The diverging health problems faced by minority groups have, in fact, been the focus of several recent initiatives of the Department of Health and Human Services, including a Task Force on Minority Health Data convened by the Public Health Service. In addition, NCHS was legislatively mandated under the Disadvantaged Minority Health Improvement Act of 1990 (P.L. 101-527) to improve the quality and quantity of information available on the health status of minority populations and subpopulations.

Race and ethnic data also are important to state and local governments to plan for schools, health and social service facilities, and many other service needs. The private sector uses these data for business planning and marketing as well as academic research. Further, race and ethnic data are an important source of information for race and ethnic groups themselves for the purposes of group identification as well as obtaining grants and services and assessing their respective social and economic statuses.

Evolution of the Race and Ethnic Questions on the 1990 Census

The increased use of race and ethnic information by government agencies and others has contributed to pressures on the Bureau to create the highest quality data possible. In addition, the race and ethnic questions in the census have been a growing source of controversy. Race and ethnicity are not objectively definable characteristics, making measurement difficult. This explains, in part, the reliance of the census on self-identification as the means for classifying respondents. In addition, various race and ethnic lobbies have pressed hard for separate status in census reporting and for more complete enumeration of their constituencies.

Because of the importance and controversial nature of race and ethnic data, the Bureau embarked on a testing and outreach program for the race and ethnic questions. According to the Bureau, improving the quality of race and ethnic statistics was one of its key objectives for the 1990 census.

1980 Census Problems Identified

The format of the race and Hispanic origin questions as they appeared in the 1980 census is shown in figure II.1.

**Appendix II
Evolution of the Race and Ethnic Questions
on the 1990 Census**

Figure II.1: 1980 Census Race and Hispanic Origin Questions

<p>4. Is this person —</p> <p><i>Fill one circle.</i></p>	<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top;"> <ul style="list-style-type: none"> <input type="radio"/> White <input type="radio"/> Black or Negro <input type="radio"/> Japanese <input type="radio"/> Chinese <input type="radio"/> Filipino <input type="radio"/> Korean <input type="radio"/> Vietnamese <input type="radio"/> Indian (Amer.) </td> <td style="width: 50%; vertical-align: top;"> <ul style="list-style-type: none"> <input type="radio"/> Asian Indian <input type="radio"/> Hawaiian <input type="radio"/> Guamanian <input type="radio"/> Samoan <input type="radio"/> Eskimo <input type="radio"/> Aleut <input type="radio"/> Other — <i>Specify</i> → </td> </tr> </table> <p><i>Print tribe</i> → _____</p>	<ul style="list-style-type: none"> <input type="radio"/> White <input type="radio"/> Black or Negro <input type="radio"/> Japanese <input type="radio"/> Chinese <input type="radio"/> Filipino <input type="radio"/> Korean <input type="radio"/> Vietnamese <input type="radio"/> Indian (Amer.) 	<ul style="list-style-type: none"> <input type="radio"/> Asian Indian <input type="radio"/> Hawaiian <input type="radio"/> Guamanian <input type="radio"/> Samoan <input type="radio"/> Eskimo <input type="radio"/> Aleut <input type="radio"/> Other — <i>Specify</i> →
<ul style="list-style-type: none"> <input type="radio"/> White <input type="radio"/> Black or Negro <input type="radio"/> Japanese <input type="radio"/> Chinese <input type="radio"/> Filipino <input type="radio"/> Korean <input type="radio"/> Vietnamese <input type="radio"/> Indian (Amer.) 	<ul style="list-style-type: none"> <input type="radio"/> Asian Indian <input type="radio"/> Hawaiian <input type="radio"/> Guamanian <input type="radio"/> Samoan <input type="radio"/> Eskimo <input type="radio"/> Aleut <input type="radio"/> Other — <i>Specify</i> → 		

<p>7. Is this person of Spanish/Hispanic origin or descent?</p> <p><i>Fill one circle.</i></p>	<ul style="list-style-type: none"> <input type="radio"/> No (not Spanish/Hispanic) <input type="radio"/> Yes, Mexican, Mexican-Amer., Chicano <input type="radio"/> Yes, Puerto Rican <input type="radio"/> Yes, Cuban <input type="radio"/> Yes, other Spanish/Hispanic
---	---

Source: 1980 Census.

In post-census evaluations, the Bureau identified several problems with the race and Hispanic origin questions from the 1980 census. In evaluating race and ethnic data, the Bureau uses several quality measures, including allocation rate and response consistency.

Allocation rate refers to a method in which a characteristic is assigned to a respondent who failed to answer the question or who failed to answer the question in a complete or legible fashion. According to Bureau staff, greater reliance on allocation procedures can lead to lower data quality.

Response consistency is a measure produced by the Content Reinterview Survey, a telephone interview or personal visit of a sample of the population, which is traditionally conducted after each decennial census.

The response a person gives to the census is compared to the response provided to the enumerator in the reinterview. Detailed probing questions are asked in the reinterview, which makes this the preferred measure. The higher the response consistency, the better the quality of data.

Three problems emerged from the Bureau's examination of the race question:

- First, the intent of the race question did not appear to be clear to some respondents, who entered ethnic groups such as German and Italian instead of marking the "White" circle. The Bureau believed this might have been a consequence of the Bureau's decision not to use the term "race" in the label for the question on the basis of some pretest results and the recommendations of some social scientists, because the question was a mixture of race and national/ethnic origin groups.
- Second, there was evidence of reporting problems through low response consistency in the "American Indian" category. Among other things, the Bureau suspected that a higher proportion of persons with some American Indian ancestry identified themselves as American Indian on the census, but when reinterviewed, they reported a different race.
- Third, there was low response consistency for "Other" in the race category. A substantial proportion of Hispanics reported "Other" for race in the census and then reported a different race when reinterviewed.

The Bureau also identified three significant problems with the Hispanic origin question:

- The first was a relatively high nonresponse rate, which required field follow-up and then the allocation of those responses that could not be filled through follow-up. The 1980 Content Reinterview Survey showed that the problem could largely be attributed to persons who were not Hispanic. These persons skipped the question when they should have indicated that they were not of Hispanic origin.
- The Hispanic origin question also suffered from misreporting by non-Hispanic Whites and Blacks who indicated they were Mexican Americans. It appears that these persons misinterpreted the "Mexican-Amer." option to mean "Mexican or American."
- Finally, the Hispanic origin question showed high inconsistency of reporting in the "Yes, other Spanish/Hispanic" category. Among other explanations, it is possible the category may have attracted non-Hispanics who wanted to indicate that they were "Other than Spanish/Hispanic."

Testing and Outreach Program Conducted

The Bureau conducted several tests of the race and Hispanic origin questions as part of the formal 1990 testing program. Race and ethnic questions were tested as part of the Bureau's census field tests in 1985, 1986, and 1987, the 1986 National Content Test, and the 1988 dress rehearsal.

In addition to the formal testing program, the Bureau conducted several special-purpose tests as part of its overall objective to improve the quality of the race and Hispanic origin data. The special-purpose tests included the 1985 Special Survey in Chicago (to gather preliminary information to design a shorter version of the race and Hispanic origin questions than appeared in the 1980 census), the 1987 Special Urban Survey (to further test versions of the race and Hispanic origin questions), the 1987 focus group sessions (to determine if terminology and instructions for race and Hispanic origin questions could be understood by members of specific race and ethnic groups), and the 1989 Special Survey. The purpose of the 1989 Special Survey was to determine if last-minute changes to the format of the race question elicited any reporting problems for which the Bureau should be prepared. These last-minute changes to the race question were the consequence of a debate between the Bureau and the Asian and Pacific Islander community discussed later in this appendix.

The Bureau also engaged in outreach efforts. These outreach efforts included two federal interagency working groups, one on race and ethnicity and one on American Indians and Alaska Natives. The interagency working groups provided an important source of information for data needs of other federal agencies, according to the Bureau. Census advisory committees, formed to represent American Indian and Alaska Natives, Asian and Pacific Islanders, Blacks, and Hispanics, provided valuable input on data needs and testing plans. The Bureau held regional meetings with officials of American Indian tribal and Alaska Native village governments to obtain suggestions on the race question.

In addition, two conferences were sponsored by the Bureau that focused on race and ethnic issues. In 1985, the Census Planning Conference on Race and Ethnic Items was held to gather input for the 1986 National Content Test. To assess the results of the National Content Test and other 1986 tests, a special meeting on race and Hispanic origin was sponsored in 1987.

Asian and Pacific Islander Concerns Dominated Debate

The most controversial change to the race and ethnic questions in the 1990 census was how Asian and Pacific Islanders were represented in the race question. The way Asian and Pacific Islanders are represented is crucial because they are the fastest growing race group in the United States, with a growth rate of 108 percent between 1980 and 1990, as figure III.4 illustrates. The changes to the race question between 1980 and 1990 are revealed by comparing the respective formats in figures 3 and II.1.

On the basis of its testing program, the Bureau recommended what is referred to as the "short version" of the race item. The short version included a single check-off item for "Asian or Pacific Islander," and replaced separate categories for Asian and Pacific Islander subgroups used in 1980 with a single space for respondents to write in their specific group. The short version of the race question, as shown in figure II.2, was included in the 1986 Los Angeles, Mississippi, and National Content Tests, among other tests. These three tests also tested longer versions of the race question, which were similar to the 1980 format. Only in these three tests were there an adequate number of persons reporting as Asian and Pacific Islander to make statistically significant comparisons between the short and longer versions of the race question, according to Bureau staff.

Figure II.2: Short Version of the Race Question

<p>4. Race</p> <p><i>Fill ONE circle for each person.</i></p> <p align="center">AND</p> <p><i>If "Asian or Pacific Islander," print one group.</i></p> <p><i>If "Indian (Amer.)," print enrolled or principal tribe</i></p> <p><i>If "Other race," print race.</i></p> <p><i>Report the race the person considers him/herself to be.</i></p>	<p>1 <input type="radio"/> White</p> <p>2 <input type="radio"/> Black or Negro</p> <p>3 <input type="radio"/> Asian or Pacific Islander <i>(Print one group, for example: Chinese, Asian Indian, Hawaiian, Laotian, Vietnamese, etc.)</i> ↗</p> <p style="border: 1px dashed black; height: 20px; margin: 5px 0;"></p> <p>4 <input type="radio"/> Indian (Amer.) <i>Print enrolled or principal tribe</i> ↗</p> <p style="border: 1px dashed black; height: 20px; margin: 5px 0;"></p> <p>5 <input type="radio"/> Eskimo</p> <p>6 <input type="radio"/> Aleut</p> <p>7 <input type="radio"/> Other race—<i>Print race</i> ↗</p> <p style="border: 1px dashed black; height: 20px; margin: 5px 0;"></p>
--	--

Note: Figure shows the question as it appeared in the 1986 Los Angeles Test. Any differences between the short version as it appeared in the Los Angeles Test and the Mississippi Test or National Content Test were minor.

Bureau Failed to Achieve Consensus Through Testing Program

Test results showed that the short version of the race question was likely to produce data on the Asian and Pacific Islander population that was as good as other test versions of the race question, which included detailed Asian and Pacific Islander categories. The Asian and Pacific Islander community, however, did not find the evidence provided by the tests strong enough to warrant such a significant change to the format of the race question. The test results were the following:

- The National Content Test and the Mississippi Test showed no statistically significant difference in the percentage of the population reporting as Asian and Pacific Islander between different versions of the race question. In the Los Angeles Test, the short version elicited a higher proportion of the population (12.6 versus 11.9 percent) reporting as Asian and Pacific Islander than a longer version of the race question with detailed Asian and Pacific Islander categories. This difference was statistically significant and was attributed to the short version capturing more persons who wrote in Asian and Pacific Islander groups such as Amerasian, Pakistani, and Sino-Vietnamese, which are not represented by the detailed categories of longer versions of the race question.

- For the most part, the percentage of the population reporting in each of the detailed Asian and Pacific Islander categories was the same between versions of the race question in the Los Angeles and National Content Tests. These detailed categories included Chinese, Japanese, Asian Indian, Hawaiian, Samoan, Korean, Guamanian, and Vietnamese. There was only one exception to this. In the National Content Test, a larger percentage of the population reported as Filipino to a version of the race question with detailed Asian and Pacific Islander categories. This difference was statistically significant. No data on reporting by detailed Asian and Pacific Islander category were available from the Mississippi Test.

The Bureau decided to use the short version in the 1990 census on the basis of these test results. Bureau staff said that, among other factors, consultations with data users also contributed to their decision to favor the short version of the race question. However, representatives of the Asian and Pacific Islander community were not persuaded. They maintained the position that the 1990 census race question should include separate categories for specific Asian and Pacific Islander groups just as the 1980 census question did.

Timely Release of Detailed Data Was Uncertain

Representatives of the Asian and Pacific Islander community were also not convinced that the Bureau could produce detailed data quickly and accurately with the short version of the race question. Although the Bureau can quickly tabulate the results of questions that use specified response categories, the short version of the race question asks for more write-in responses, which are more difficult to process. In 1980, census write-in data had to be manually keyed and then tabulated through a clerical review process. Delays in the publication of detailed Asian and Pacific Islander data after the 1980 census resulted in concerns about how the data from the 1990 census would be processed.

The Bureau promised to undertake an automated coding operation in processing 1990 census data to ensure that write-in responses to the race question could be tabulated by computer for all questionnaires. This means that write-in responses, once manually keyed, are matched by computer against a master code list. A Bureau official said that this promise of automated coding may have come too late to persuade representatives of the Asian and Pacific Islander community, where opposition to the short version had grown strong. Bureau staff said that the Bureau was late in responding to these concerns because the

technology for undertaking such an operation was not developed until late in the decade.

Bureau Reversed Its Position

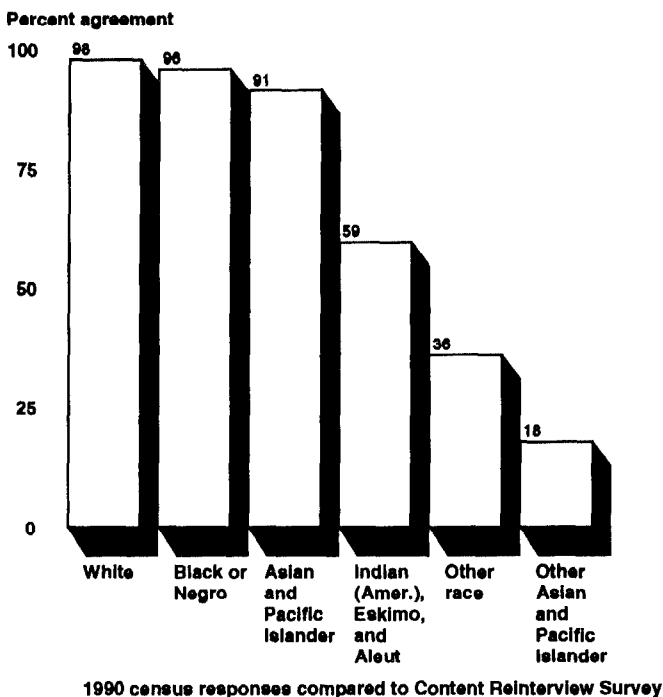
The Asian and Pacific Islander community succeeded in gaining congressional attention for its concerns. The conference report accompanying the appropriations bill for the Commerce Department for fiscal year 1989 contained language indicating that the Census Bureau should use prelisted categories for Asian and Pacific Islander groups as in 1980. Another bill requiring detailed categories and tabulations of the Asian and Pacific Islander population for the 1990 census and subsequent censuses was pocket vetoed by President Reagan. Responding to congressional direction and pressures from the Asian and Pacific Islander community, the Bureau reconsidered its original decision and chose to include in the 1990 census a version of the race question with prelisted Asian and Pacific Islander categories.

Quality of Race and Ethnic Data

Indicators Are Generally Positive

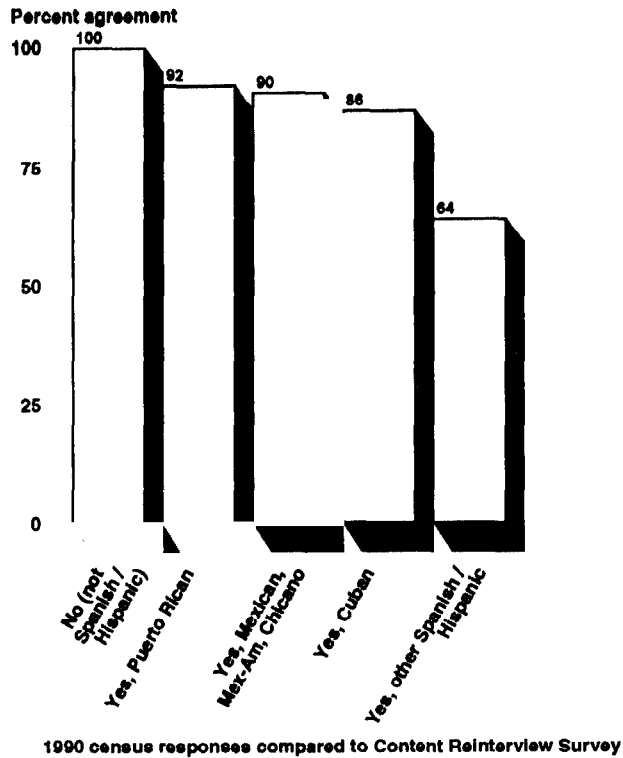
The Content Reinterview Survey for the 1990 census showed generally good response consistency for both the race and Hispanic origin questions, i.e., persons reported the same race and origin in the reinterview as they did in the census. The percentage of agreement between responses to the census and the reinterviews is illustrated in figures III.1 and III.2. There are a few categories, however, that showed reporting problems.

Figure III.1: Response Consistency for Race Question



Source: Preliminary data provided by the Census Bureau.

Figure III.2: Response Consistency for Hispanic Origin Question



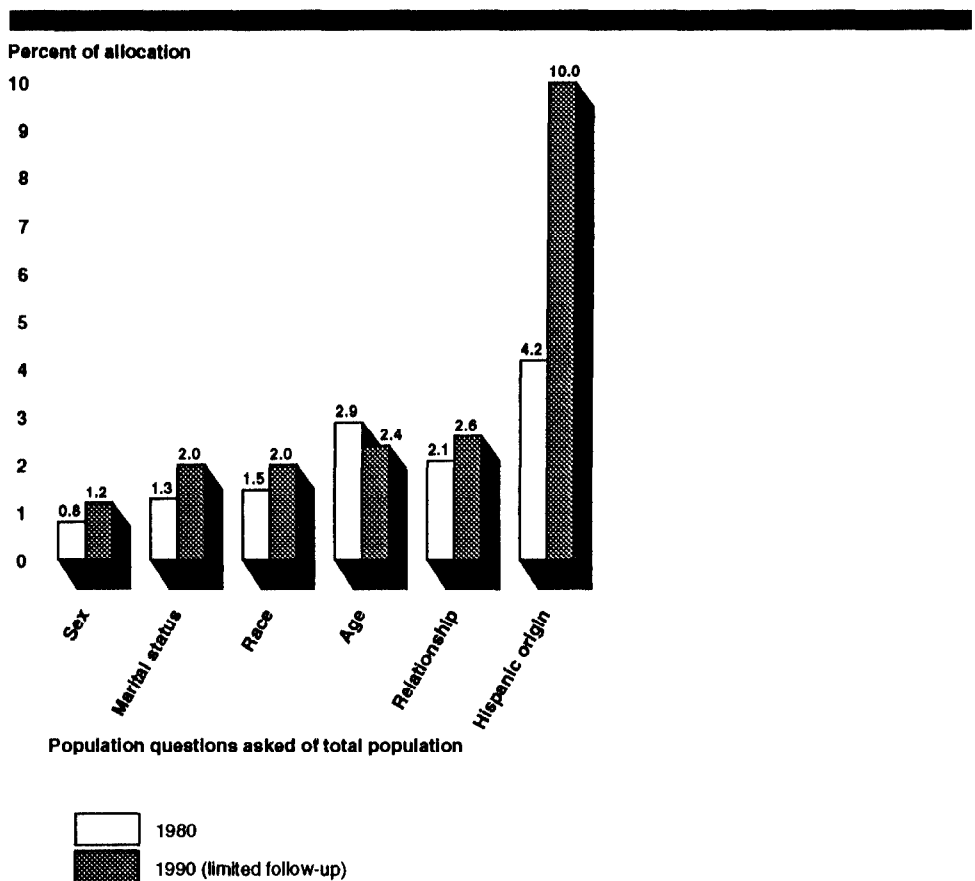
Source: Preliminary data provided by the Census Bureau.

Three of the six race categories showed lower response consistency, as represented in figure III.1. These categories are “Indian (Amer.)/ Eskimo/Aleut,” “Other race,” and “Other Asian or Pacific Islander.” Much of the inconsistent reporting in the race question was attributed by the Bureau to Hispanics, particularly foreign-born Hispanics, who have difficulty classifying themselves by race. As illustrated in figure III.2, the only category in the Hispanic origin question in which the Bureau found response consistency to be relatively low was the “Yes, other Spanish/Hispanic” category. Overall, these results showed that response consistency problems experienced in 1980 persisted in 1990.

According to the Bureau, another important measure of quality is the allocation rate. The 1990 allocation rate for the race question was

2 percent. As shown in figure III.3, this allocation rate is comparable to the allocation rates of other questions that were asked of the total population and that are subject to much less controversy—with the exception of Hispanic origin, for which the allocation rate was much higher. This rate, however, represents an increase from the 1.5 percent allocation rate in the 1980 census.

Figure III.3: Comparative Allocation Rates



Source: Census Bureau.

In addition, the results from the 1990 census showed that the Hispanic origin item continues to pose one of the more significant data quality challenges for the Bureau in terms of allocation rate. The 1990 Hispanic

item had an allocation rate of 10 percent. This was the highest allocation rate of population questions asked of all persons, as shown in figure III.3.

The considerable difference between allocation rates for the Hispanic origin question between the 1980 and 1990 censuses is attributed for the most part by the Bureau to a lower level of follow-up in 1990. Follow-up includes a telephone call or a personal visit by a census enumerator when a questionnaire has inconsistent or missing population data beyond an acceptable level. This decision to reduce follow-up in the 1990 census was made to accommodate unanticipated cost increases and to fund new program priorities. When follow-up is reduced, missing or inconsistent data is filled in by allocation procedures. A higher allocation rate may not be problematic if Hispanic origin characteristics are allocated in an unbiased fashion. Preliminary analysis showed no evidence of bias, but Bureau staff said that further study is needed before any conclusions about this or other impacts can be drawn.

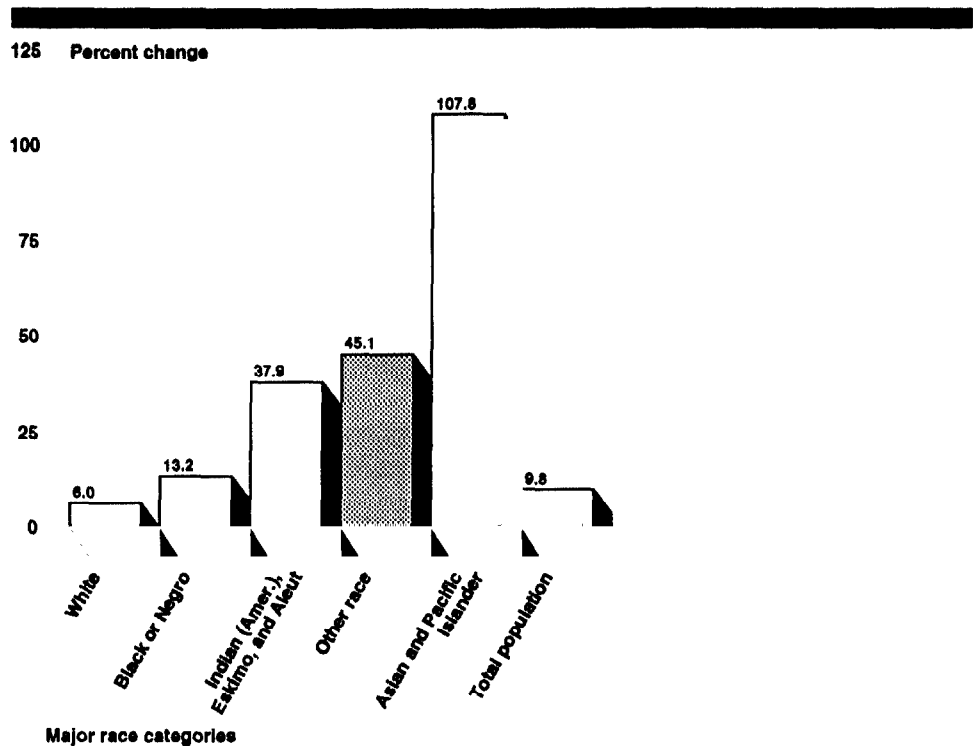
The Bureau has identified several underlying problems that contributed to the higher allocation rate for the Hispanic origin question in both the 1980 and 1990 census as compared to other questions. There is a high rate of nonresponse among persons who are not Hispanic and simply do not see the question as relevant to them. In addition, confusion Hispanics experience with the race item may spill over to problems with the Hispanic origin item. Some Hispanics equate their "Hispanicity" with race by responding "Other race" in the race item, indicating they are Hispanic in the space the race item provides, and then skipping over the Hispanic origin item because they see this item as superfluous.

The "Other Race" Category Is Growing

Almost 10 million people were reported in the "Other race" category. A greater percentage of respondents—almost 4 percent—were reported in the "Other race" category in the 1990 census as compared to the last several censuses.

Although this is still a relatively small percentage of the total U.S. population, the "Other race" category has grown at a fast rate for the last couple of decades. Between 1980 and 1990, the "Other race" category grew by 45 percent (from approximately 7 million to 10 million). This is a faster growth rate than any other race group, with the exception of Asian and Pacific Islanders, as shown in figure III.4. This growth rate is much faster than the total population growth rate of less than 10 percent.

Figure III.4: Growth of Race Groups
From 1980 to 1990



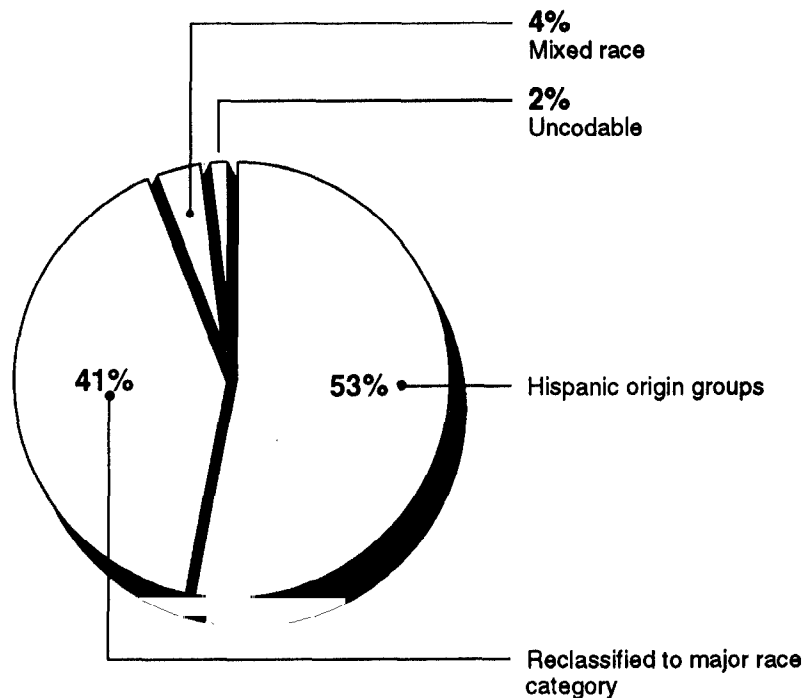
Source: Census Bureau.

Most "Other Race" Respondents Are Hispanic

According to the Bureau, 97.5 percent of the almost 10 million persons who are reported in the "Other race" category in the 1990 census are persons of Hispanic origin. This is consistent with the 1980 census in which the Bureau estimates 95 percent of persons reported in the "Other race" category were Hispanic. The Bureau believes this reflects the difficulties the race item poses for some Hispanics.

Although only a portion of respondents who mark the "Other race" category provide write-in responses, the Bureau has examined these responses to help identify the causes of reporting problems. Characteristics of the approximately 2.5 million write-ins to the "Other race" category are illustrated in figure III.5.

Figure III.5: Write-in Responses to
"Other Race" Category



Source: Census Bureau.

Forty-one percent of the write-ins were reclassified through automated coding to another race category (e.g., White, Black, American Indian, Asian or Pacific Islander) and, as such, were not included in the final counts for "Other race." These persons had indicated an ethnic group such as Irish, Arab, or Haitian. This means that more than 1 million non-Hispanic respondents were not able to, or chose not to, assign themselves to a major race category provided. The bulk of the remaining persons were of Hispanic origin, which is consistent with the Bureau's overall findings that most persons reported in the "Other race" category are Hispanic.

About 4 percent of those persons who provided a write-in response to the "Other race" category specified "mixed" or "biracial." Although this is a relatively small portion, many more would identify with mixed race if the option were provided, according to Bureau demographers. Evidence of

this includes the fact that some respondents check two or more race categories. The computer, however, reads only the response that is marked the darkest, and the respondent is identified as only one race. Others might write in a response such as "Black-White," which is recorded as "Black" because it is the first of the two identifications. Only 2 percent of the responses were uncodable responses.

Except for Hispanics, the Bureau has experienced relative success in classifying persons to major race categories and minimizing the number of persons assigned to the "Other race" category. There are indications, however, of the growing diversity in the population. To capture the 8 million write-ins to the race question (which include write-ins to "Other race," "Indian (Amer.)," and "Other API"), the Bureau added over 200,000 codes in the automated coding process. For example, persons who wrote "German" and marked the "Other race" category were assigned a detailed code by the Bureau so as to capture the write-in information and then were reclassified in the "White" category. The Bureau has cited several factors as possibly contributing to the addition of so many new codes. These factors include population migration to the United States over the last decade as well as the coding of all write-in data to the race question for the first time in census history.

Bureau Reclassified Persons in "Other Race" Category

Because other data collection systems do not include an "Other race" category, the Bureau creates a "modified race file." To create this file, the Bureau redistributes persons in the "Other race" category to one of the race groups accepted by Directive No. 15 to make census data compatible with race data from other collection systems. This special data file is designed to meet the needs of federal and state agencies and researchers.

However, the Bureau is challenged with developing a method for allocating a race to those in the "Other race" category. Persons in the "Other race" category were assigned the same race of the nearest person who gave an identical response to the Hispanic origin question. Hispanic origin figures heavily in the methodology adopted because the vast majority of those who report in the "Other race" category are of Hispanic origin. However, some demographers recognize that the Bureau's choice of methodology is sensitive because different race assignment rules to construct the "modified race file" may produce different counts for major race groups.

Status of Census Bureau Research on Race and Ethnicity

Indications are that measuring race and ethnicity will continue to challenge the Bureau in the 2000 census and beyond. The Bureau anticipates that, as minority populations grow, the Bureau will be subject to increasing pressure from new groups and subgroups for identification on the census form and in census data products. Among those lobbying the Bureau for new designations in the year 2000 are groups seeking a biracial or multiethnic category for children of intermarriages. At the same time, the Bureau has the overall goal of simplifying the questionnaire in order to improve the census response rate.¹

Bureau Is Developing Strategy for Research and Development

The Bureau has a plan for research and development of race and ethnic questions for the 2000 census. The plan lays out the basic steps that will be part of the Bureau's research and development effort, including consulting with race and ethnic groups and testing race and ethnic questions.

Bureau staff said that the Bureau hopes to commit more time to evaluating the 1990 census so the agency is better positioned to focus on results in planning for the 2000 census. Pressures for getting the 1990 race and ethnic data products issued, however, have constrained the Bureau's ability to do evaluation. Bureau staff indicated that these competing pressures may lead to some delays in testing and consultation with key customers, including representatives of race and ethnic groups.

The Bureau's first field tests of the race and ethnic questions were originally scheduled for fiscal year 1993, but in the latest version of the plan they are slated for fiscal year 1994. Bureau officials said they expect the first tests to be targeted to particular ethnic groups. Then, in the 1995 census test, the Bureau will have the opportunity to test a version of the race and ethnic questions that shows the greatest promise on the basis of research to date. The 1995 census test, as a field test of several census design options, is an important component of the Bureau's 2000 census planning effort. However, the Bureau is faced with the prospect that the 1994 test results may not be available in time to incorporate in the 1995 test.

¹In *Census Reform: Questionnaire Test Shows Simplification Holds Promise* (GAO/T-GGD-92-59, July 1, 1992) we discussed the preliminary results and implications of the Bureau's test of simplified census questionnaires. The test showed that a shorter questionnaire can lead to better response rates.

Bureau Plans to Expand Consultation Program

A major component of the Bureau's research and development plan is a consultation program. The consultation program for the 2000 census is intended to build communication between the Bureau and its customers, including organizations representing the various race and ethnic groups. The Bureau also plans to better coordinate its consultation program with its research and testing effort. The Bureau's goal is for communication to be developed with its customers earlier in the process than was the case in the 1990 census so that last-minute changes in content can be avoided.

The importance of coming to a broad agreement with race and ethnic groups early in the process is evidenced in the controversy regarding how the Asian and Pacific Islander groups would be represented in the race question. In congressional hearings, representatives of the Asian and Pacific Islander community said that the Bureau had not solicited their participation in the early phases of redesigning the race question. Representatives to the census advisory committees for minority populations made a similar point in discussions with us. Several representatives said that the Bureau had already formulated the census questions before the committees began to meet.

The advisory committees were established in 1986 for the 1990 census by the Department of Commerce. This was the same year that the major tests were held that drove the debate on the race question. Bureau staff agreed that the formation of these committees so late in the planning process contributed to problems in gaining consensus. Even so, the Bureau may be faced with the same obstacle in planning for the 2000 census because the charters for the committees for the 1990 census are expected to expire at the end of fiscal year 1993, and no date for the formation of new committees for the 2000 census has been established.

Bureau staff said that its consultation effort for the 2000 census will be broader and more intensive, and it will involve new and emerging ethnic groups in addition to the ethnic groups now represented in the Bureau's outreach efforts. The Bureau has developed a consultation plan that identifies the groups that are to be involved and the schedule of consultations. Bureau staff said that they hope the census advisory committees for minority populations will be reconstituted soon as part of the Bureau's overall consultation effort, and they have set fiscal year 1994 as the target for beginning consultations with these committees.

Bureau Is Issuing More Data Products and Estimates

Race and Hispanic origin data from the decennial census generally met the needs of the federal agencies we contacted. Compared to the 1980 census, the Bureau issued more census data products on race and ethnicity after the 1990 census on a more timely basis. Among the new products, the Bureau issued state profiles for race and Hispanic origin as well as data for cities and metropolitan areas. New data processing technology, including automated coding, has been credited for enabling the Bureau to improve upon its performance after the 1980 census.

However, a changing population and growing program needs will put pressure on the Bureau to develop ever more current and detailed data, which is beyond the scope of the decennial census. One problem staff from the agencies we contacted identified was the aging of census data. Intercensal estimates, which serve to update the census, were not detailed enough to meet the needs of these agencies, nor were they available for smaller geographic areas.

For example, a Department of Health and Human Services initiative entitled "Healthy People 2000" has created a greater need for data. The National Center for Health Statistics is responsible for tracking a wide range of health objectives by race and ethnic group at the national, state, and local levels. To carry out this responsibility, NCHS needs more current and detailed data than the census provides. In addition, Equal Employment Opportunity Commission staff said that the lack of current data has had an impact on their efforts to advance equal opportunity for minorities. Several judges have disallowed evidence EEOC has provided to demonstrate patterns of discrimination in certain localities on the basis that the data were not up-to-date.

To try to meet the growing demand for race and ethnic data, the Bureau is expanding its estimates program. For example, national estimates for this decade will grow from three race groups (White, Black, Other) to four race groups (White, Black, Asian and Pacific Islander, and American Indian and Alaska Native) by ethnic origin (Hispanic, non-Hispanic). The Bureau has also launched several experimental programs designed to provide estimates of race and ethnic groups for some areas below the national level.

While the demand for more detailed estimates is increasing, Bureau staff said population estimates for race and ethnic groups are becoming more difficult to produce. Bureau staff and other experts cite many different obstacles to developing estimates. A primary obstacle is that most

**Appendix IV
Status of Census Bureau Research on Race
and Ethnicity**

administrative records capture less race and ethnic detail than the census. This is problematic because the Bureau relies heavily on administrative records to develop estimates. The Bureau is thus faced with demands for detailed race and ethnic data that not only extend beyond the scope of the decennial census but also challenge its capacities to produce reliable estimates.

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are \$2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

**U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015**

or visit:

**Room 1000
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC**

**Orders may also be placed by calling (202) 512-6000
or by using fax number (301) 258-4066.**

**United States
General Accounting Office
Washington, D.C. 20548**

**Official Business
Penalty for Private Use \$300**

**First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100**
