

Symposium: Race/Ethnicity and the 2000 Census

Measuring Our Nation's Diversity: Developing a Common Language for Data on Race/Ethnicity

Katherine K. Wallman, BA, Suzann Evinger, MA, and Susan Schechter, MA

ABSTRACT

During the 4-year period 1993 through 1997, the Office of Management and Budget (OMB) undertook a comprehensive review of the statistical standards that are used throughout the federal government to gather and publish data on race and ethnicity. The primary objective of this review was to ensure that our standards provide a common language that reflects the increasing diversity of the US population and maintains our ability to monitor compliance with civil rights laws. The review culminated with the October 1997 issuance of OMB's "Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity." In this article, we describe key aspects of the process that was undertaken to review and revise the 1977 standards. We also attempt to dispel some myths and misunderstandings that have been associated with these standards. (*Am J Public Health*. 2000;90:1704-1708)

Perhaps the greatest myth associated with the federal government's collection of data on race and ethnicity is that comparable data have been available over the years within and across agencies. One need only explore Margo Anderson's excellent volume *The American Census: A Social History* to realize that since 1790 we have been collecting—and changing each decade the way we collect—the data on race and ethnicity that provide denominators for much of our social and demographic analysis.¹

If we "fast-forward" to the mid-1970s, we find ourselves at the beginning of the federal government's initial efforts to develop and implement a common language for reporting data on race and ethnicity. By that time—largely but not exclusively as a consequence of the civil rights laws of the 1960s, the Voting Rights Act, and a small (as measured by number of words) law known as Public Law 94-311 that required all federal agencies to provide separate counts of the Hispanic population in data collections—agencies within and across departments increasingly were collecting data related to race and ethnicity. Those agencies, led in particular by the Office of Education and its parent, the Department of Health, Education, and Welfare, came to the Office of Management and Budget (OMB) seeking our leadership in developing government-wide standards for data on race and ethnicity.

And thus was born—after an interagency effort of more than 3 years—the first federal government standard for the collection and presentation of such data. The standard specified a minimum set of categories, including 4 categories for data on race (American Indian or Alaskan Native, Asian or Pacific Islander, Black, and White) and 2 categories for data on ethnicity ("Hispanic origin" and "Not of Hispanic origin"). The birth name of this standard, long since forgotten, was OMB Circular No. A-46, Exhibit F. It experienced a legal name change in 1978, when it was dubbed "Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting."

By the early 1990s, when signals from within and outside the federal government suggested that the standard might be failing to keep pace with changes in our nation's population, the "common language" for reporting data on race and ethnicity had in fact been in place for only 15 years. The increasingly frequent expressions of concern were motivated principally by 2 types of encounters that bring the public face to face with our data standards. The first of these was the 1990 decennial census. Many constituents complained not only to the Census Bureau and OMB but also to members of Congress that they could not find themselves in the choices offered. Particularly distressed by the requirement to indicate a single race, some half a million individuals selected multiple races in the 1990 census. The second type of encounter that evoked substantial concern was the experience of parents registering children from the growing number of interracial unions for school and being asked, as they so often told us, to "select one parent and deny the other."

Motivated by the demographic changes taking place and the growing expressions of concern about the relevance of current requirements for data on race and ethnicity, OMB agreed in 1993 to undertake a review of the 1977 standards, with no prejudgment that any change would be made.²⁻⁴ The challenges we faced were complex. Simply put, we were asked to develop a system to classify data on race and ethnicity that would capture the nation's increasing diversity and, simultaneously, maintain our ability to measure and track social and economic conditions for populations that historically had experienced discrimination.

The authors are with the United States Office of Management and Budget, Washington, DC.

Requests for reprints should be sent to Katherine K. Wallman, United States Office of Management and Budget, 725 17th St NW, Room 10201, Washington, DC 20503 (e-mail: kwallman@omb.eop.gov).

This article was accepted June 25, 2000.

The Review Process

The process we undertook was also complex. It involved a considerable program of research, a 30-agency committee, and numerous opportunities for public comment that included public hearings around the country (various phases of the review have been addressed in several *Federal Register* notices⁵⁻⁸).⁹ The public comment process was especially useful in identifying a number of issues that needed to be addressed, including the following:

- How should data on individuals of multiple racial heritages be classified?
- Should “Hispanic” be a response option to the race question?
- Should data on race and ethnicity be gathered via 2 separate questions? If yes, then what should be the sequence of these questions? Or should data on race and ethnicity be gathered in a single question?
- Should data on Native Hawaiians continue to be classified in the Asian or Pacific Islander category?
- Should the minimum set of categories for data on race and ethnicity be expanded to include other population groups?
- Should the terminology for some of the categories be changed?

The program for research and testing that was developed under the auspices of the interagency committee addressed significant issues that had been identified in the public comments. Three major sample surveys were conducted: the May 1995 Current Population Survey (CPS) Supplement on Race and Ethnicity, the 1996 National Content Survey, and the 1996 Race and Ethnic Targeted Test. The CPS supplement was designed to test the effects of separate questions on race and ethnicity, with or without a “multiracial” category, and the effects of a combined question on race and ethnicity, with or without a “multiracial” category.¹⁰ The National Content Survey was designed to evaluate the effects of adding a “multiracial or biracial” category and the effects of having the Hispanic origin question asked before or after the race question.¹¹ The Race and Ethnic Targeted Test was designed to test the effects of (1) a “multiracial or biracial” category, (2) “mark one or more” and “mark all that apply” approaches to reporting more than 1 race, and (3) a combined question on race and Hispanic origin.¹²

In addition, the National Center for Education Statistics and the Office for Civil Rights of the Department of Education jointly conducted a survey to determine how schools collect data on the race and ethnicity of their students and how the administrative records containing these data are maintained to meet statutory requirements for reporting aggregate information to the federal government.¹³ Be-

cause the mode of data collection can affect responses, the research agenda proposed studies involving both in-person or telephone interviews and self-administered questionnaires, such as the decennial census, that are completed by the respondent and mailed back. Cognitive interviews were conducted with various groups to provide guidance on the wording of the questions and the instructions for the tests and studies.

Key Issues and Recommendations for Change

The most controversial and sensitive issue—and the one that continues to receive the most media attention and generate the most public discussion—concerned whether and how the standards should be changed to classify data on individuals who have parents of different races and who wish to identify with more than 1 race. Sample surveys were designed to test the effects of a category labeled “multiracial” and to test alternative approaches for permitting respondents to select more than 1 race. The research conducted under the auspices of the interagency committee indicated that less than 2% of the population selected 2 or more races when given the opportunity to do so but that the proportion of the population identifying with multiple racial heritages seemed to be growing. Furthermore, research indicated that a considerable number of respondents were selecting more than 1 racial category even when instructed to choose only one and that there was not a shared understanding of the term *multiracial*. The interagency committee recommended, and OMB agreed, that

- The standard should allow individuals who so desire to reflect more than 1 race when responding to federal data collections.
- There should not be a separate racial category (a check box) labeled “multiracial.”
- When data are reported, a minimum of 1 additional racial category, designated “more than one race,” must be included, if the criteria for data quality and confidentiality are met, to allow reporting of the aggregate number of multiple race responses. Data producers are encouraged to provide greater detail about the distribution of multiple responses.

The interagency committee also conducted research on the issue of whether there should be a combined race/Hispanic origin question instead of 2 separate questions on race and Hispanic origin. Research indicated that in the 1990 census—which involved separate questions, the race question preceding the Hispanic origin question—a large percentage of Hispanics had selected “other race.” As a result, the committee tested separate questions and their sequencing as compared with a combined

question. This research was designed to determine which format produced more complete data on Hispanics and non-Hispanics, that is, which format worked best in terms of both reducing nonresponse to the Hispanic origin question and reducing Hispanics’ selection of “other race” as a response to the race question. The interagency committee recommended, and OMB agreed, that the Hispanic origin question should be asked separately and should be placed before the race question.

Public comment and testimony of witnesses at OMB and congressional hearings included a variety of requests to expand the minimum set of categories by establishing additional categories for specific population groups, including Middle Easterners/Arabs, Cape Verdeans, European Americans, German Americans, and Creoles. Although it has been required that the minimum set of categories established in 1977 be used in all federal data collections, it is important to note that many federal data collections, including the decennial census, request considerably more detailed data on race and ethnicity.

Moreover, the decennial census includes an “ancestry” question on the “long form” that collects additional information on population groups such as those requesting the expansion of the minimum set of categories in the government-wide standard. The review considered, among other matters, legislative needs, definitional issues, the size and geographic concentration of several of these populations, and the feasibility of gathering representative information in all federal data collections. The interagency committee recommended, and OMB agreed, that there should be no racial or ethnic categories added to the 1977 minimum standard.

The interagency committee also reviewed requests to remove Native Hawaiians from the Asian or Pacific Islander category. It considered requests to reclassify Native Hawaiians in the American Indian or Alaska Native category so that all indigenous peoples would be included in the same category, as well as requests to classify Native Hawaiians in a separate category. In particular, the review was concerned with the effects a reclassification would have on the usefulness of the social and economic data for the respective population groups. Whereas Hawaiians represent almost 3% of the Asian or Pacific Islander population, they would constitute almost 10% of a more inclusive category for indigenous groups.⁷ OMB ultimately concluded that the Asian or Pacific Islander category should be divided into 2 categories, “Asian” and “Native Hawaiian or Other Pacific Islander.”

Finally, the research program included testing of terminology changes that had been suggested in public comment. This testing was useful in determining respondents’ preferences for terminology, their understanding of the meaning of the terms, and whether—and in

This classification provides a minimum standard for maintaining, collecting, and presenting data on race and ethnicity for all Federal reporting purposes. The categories in this classification are social-political constructs and should not be interpreted as being scientific or anthropological in nature. They are not to be used as determinants of eligibility for participation in any Federal program. The standards have been developed to provide a common language for uniformity and comparability in the collection and use of data on race and ethnicity by Federal agencies. The standards have five categories for data on race: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. There are two categories for data on ethnicity: "Hispanic or Latino," and "Not Hispanic or Latino."

1. Categories and Definitions

The minimum categories for data on race and ethnicity for Federal statistics, program administrative reporting, and civil rights compliance reporting are defined as follows:

- *American Indian or Alaska Native.* A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.
- *Asian.* A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
- *Black or African American.* A person having origins in any of the black racial groups of Africa. Terms such as "Haitian" or "Negro" can be used in addition to "Black or African American."
- *Hispanic or Latino.* A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. The term, "Spanish origin," can be used in addition to "Hispanic or Latino."
- *Native Hawaiian or Other Pacific Islander.* A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
- *White.* A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

Respondents shall be offered the option of selecting one or more racial designations. Recommended forms for the instruction accompanying the multiple response question are "Mark one or more" and "Select one or more."

1. Data Formats

The standards provide two formats that may be used for data on race and ethnicity. Self-reporting or self-identification using two separate questions is the preferred method for collecting data on race and ethnicity. In situations where self-reporting is not practicable or feasible, the combined format may be used.

In no case shall the provisions of the standards be construed to limit the collection of data to the categories described above. The collection of greater detail is encouraged; however, any collection that uses more detail shall be organized in such a way that the additional categories can be aggregated into these minimum categories for data on race and ethnicity. With respect to tabulation, the procedures used by Federal agencies shall result in the production of as much detailed information on race and ethnicity as possible. However, Federal agencies shall not present data on detailed categories if doing so would compromise data quality or confidentiality standards.

- a. *Two-question format.* To provide flexibility and ensure data quality, separate questions shall be used wherever feasible for reporting race and ethnicity. When race and ethnicity are collected separately, ethnicity shall be collected first. If race and ethnicity are collected separately, the minimum designations are:

Race

- American Indian or Alaska Native
- Asian
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White

Ethnicity

- Hispanic or Latino
- Not Hispanic or Latino

When data on race and ethnicity are collected separately, provision shall be made to report the number of respondents in each racial category who are Hispanic or Latino. When aggregate data are presented, data producers shall provide the number of respondents who marked (or selected) only one category, separately for each of the five racial categories. In addition to these numbers, data producers are strongly encouraged to provide the detailed distributions, including all possible combinations, of multiple responses to the race question. If data on multiple responses are collapsed, at a minimum the total number of respondents reporting

"more than one race" shall be made available.

- b. *Combined format.* The combined format may be used, if necessary, for observer-collected data on race and ethnicity. Both race (including multiple responses) and ethnicity shall be collected when appropriate and feasible, although the selection of one category in the combined format is acceptable. If a combined format is used, there are six minimum categories:
 - American Indian or Alaska Native
 - Asian
 - Black or African American
 - Hispanic or Latino
 - Native Hawaiian or Other Pacific Islander
 - White

When aggregate data are presented, data producers shall provide the number of respondents who marked (or selected) only one category, separately for each of the six categories. In addition to these numbers, data producers are strongly encouraged to provide the detailed distributions, including all possible combinations, of multiple responses. In cases where data on multiple responses are collapsed, the total number of respondents reporting "Hispanic or Latino and one or more races" and the total number of respondents reporting "more than one race" (regardless of ethnicity) shall be provided.

2. Use of the Standards for Record Keeping and Reporting

The minimum standard categories shall be used for reporting as follows:

- a. *Statistical reporting.* These standards shall be used at a minimum for all federally sponsored statistical data collections that include data on race and/or ethnicity, except when the collection involves a sample of such size that the data on the smaller categories would be unreliable, or when the collection effort focuses on a specific racial or ethnic group. Any other variation will have to be specifically authorized by the Office of Management and Budget (OMB) through the information collection clearance process. In those cases where the data collection is not subject to the information collection clearance process, a direct request for a variance shall be made to OMB.
- b. *General program administrative and grant reporting.* These standards shall be used for all Federal administrative reporting or record keeping requirements that include data on race and

Continued

Source. Office of Management and Budget. Revisions to the standards for the classification of federal data on race and ethnicity (October 30, 1997). Standards for maintaining, collecting, and presenting federal data on race and ethnicity. Available at: <http://www.whitehouse.gov/omb/fedreg/ombdir15.html>. Accessed September 7, 2000.

FIGURE 1—Standards for maintaining, collecting, and presenting federal data on race and ethnicity.

ethnicity. Agencies that cannot follow these standards must request a variance from OMB. Variances will be considered if the agency can demonstrate that it is not reasonable for the primary reporter to determine racial or ethnic background in terms of the specified categories, that determination of racial or ethnic background is not critical to the administration of the program in question, or that the specific program is directed to only one or a limited number of racial or ethnic groups.

c. *Civil rights and other compliance reporting.* These standards shall be used by all Federal agencies in either the separate or combined format for civil rights and other compliance reporting from the public and private sectors and all levels of government. Any variation requiring less detailed data or data which cannot be aggregated into the basic categories must be specifically approved by OMB for executive agencies. More detailed reporting which can be aggregated to the basic categories may be used at the agencies' discretion.

3. Presentation of Data on Race and Ethnicity

Displays of statistical, administrative, and compliance data on race and ethnicity shall use the categories listed above. The term "nonwhite" is not acceptable for use in the presentation of Federal Government data. It shall not be used in any publication or in the text of any report.

In cases where the standard categories are considered inappropriate for presentation of data on particular programs or for particular regional areas, the sponsoring agency may use:

- a. The designations "Black or African American and Other Races" or "All Other Races" as collective descriptions of minority races when the most summary distinction between the majority and minority races is appropriate;
- b. The designations "White," "Black or African American," and "All Other Races" when the distinction among the majority race, the principal minority race, and other races is appropriate; or
- c. The designation of a particular minority race or races, and the inclusion of "Whites" with "All Other

Races" when such a collective description is appropriate.

In displaying detailed information that represents a combination of race and ethnicity, the description of the data being displayed shall clearly indicate that both bases of classification are being used.

When the primary focus of a report is on two or more specific identifiable groups in the population, one or more of which is racial or ethnic, it is acceptable to display data for each of the particular groups separately and to describe data relating to the remainder of the population by an appropriate collective description.

4. Effective Date

The provisions of these standards are effective immediately for all new and revised record keeping or reporting requirements that include racial and/or ethnic information. All existing record keeping or reporting requirements shall be made consistent with these standards at the time they are submitted for extension, or not later than January 1, 2003.

FIGURE 1—Continued.

what ways—any proposed changes in terminology might affect reporting. OMB ultimately decided that (1) the Black category should be changed to "Black or African American" (additional terms, such as Haitian or Negro, can also be used), (2) the term "Hispanic" should be changed to "Hispanic or Latino" (terms such as "Spanish origin" can also be used), (3) the term "American Indian" should not be changed to "Native American," and (4) the term "Hawaiian" should be changed to "Native Hawaiian."

The revised "Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity" were issued by OMB on October 30, 1997 (Figure 1).

Some Caveats Concerning the Standards

In reviewing OMB's decisions on revising the standards, it is important to recall that the categories for data on race and ethnicity are a product of US political and social history. Moreover, as stated in both Directive 15 and the October 1997 revised standards, the categories should not be interpreted as having anthropological or scientific origins. The driving force for the development of the standards in the 1970s was the need for comparable data to monitor equal access, in areas such as housing, education, mortgage lending, health care services, and employment opportunities, for population groups that historically had experi-

enced discrimination and differential treatment because of their race or ethnicity. When standard categories initially were adopted by OMB in May 1977, they reflected in many respects legislatively based priorities for data on particular population groups.

The 1997 OMB decisions retain the concept that the standards provide a *minimum* set of categories for use throughout the federal government for all record keeping, collection, and presentation of comparable data on race and ethnicity. This minimum set of categories is used not only in the decennial census (which provides the "denominator" for many measures) but also in all federal data collections (e.g., those related to medical research). Consequently, expanding the minimum set of categories would have implications for all federal surveys and forms, as well as for the business community, state and local governments, and other private sector organizations that are required to use the minimum set to meet federal record-keeping and reporting requirements.

Furthermore, it should be noted that the OMB decisions do not identify or designate certain population groups as "minority groups." The decisions continue the policy that the categories are not to be used for determining eligibility of population groups for participation in any federal programs. The standards do not contain criteria or qualifications (such as blood quantum levels) that are to be used in determining a particular individual's racial or ethnic

classification. The standards do not tell an individual who he or she is or specify how an individual should classify himself or herself; self-identification continues to be the preferred means of obtaining data on an individual's race and ethnicity.

Implementation Challenges: A Work in Progress

In announcing the revised standards, OMB recognized the need for further research, particularly with respect to tabulation of data. Issues identified for such research included the following:

- How should data be used to evaluate conformance with program objectives in the area of equal employment opportunity and other antidiscrimination programs?
- How should the decennial census data for many small population groups with multiple racial heritages be used to develop sample designs and survey controls for major demographic surveys?
- How should the revised standards be introduced in the vital statistics program, which obtains the number of births or deaths from administrative records but uses intercensal population estimates in determining rates of births and deaths?
- More generally, how can meaningful comparisons be made of data collected under

the earlier standards and data that will be collected under the revised standards?

These and other issues concerning the collection, tabulation, and bridging of data on race and ethnicity have been and continue to be explored by interagency teams.

Initial efforts to address these issues were described in a report released by OMB in February 1999, *Draft Provisional Guidance on the Implementation of the 1997 Standards for the Collection of Federal Data on Race and Ethnicity*.¹⁴ The report provided guidance focused on 3 areas: collecting data under the revised standards, tabulating data collected under the revised standards, and building bridges to compare data collected under the revised and earlier standards. Since the issuance of the draft provisional guidance report, additional research and analyses have been completed, and discussions with stakeholders within and outside the federal government have further informed various proposals. Early work addressed tabulation plans for the 2000 census data products. In addition, in response to requests from agencies responsible for monitoring and enforcing civil rights laws, OMB led an interagency group formed to develop guidance on (1) the collection of aggregate data on race when agencies request information from businesses, schools, and other entities and (2) the allocation by agencies of multiple race responses, whether individual or aggregate, for use in civil rights monitoring and enforcement. This guidance (issued as OMB Bulletin 00-02, dated March 9, 2000¹⁵) ensures that agencies can continue to monitor compliance with laws that offer protections for those who historically have experienced discrimination and that reporting burdens are minimized for those reporting aggregate data to federal agencies.

For some aspects of the implementation guidance, work is ongoing; the guidance will be amended as additional research and analy-

ses are completed. Late in 2000, OMB expects to issue for public comment a revised version of the provisional guidance. OMB's guidance will continue to evolve as data from the 2000 census and other information collections employing the 1997 standards become available.

The review and ultimate revision of OMB's standards for data on race and ethnicity was a complex task that actively involved the policy and research communities as well as the general public. With the issuance of the revised standards in 1997, OMB reached a major milestone. But many challenges remain. □

Contributors

K. K. Wallman directed the review of the standards and participated in the writing and editing of the paper. S. Evinger was the principal staff member on the project, wrote major portions of the paper, and contributed to the editing. S. Schechter provided additional materials about recent developments and contributed to the revisions.

References

1. Anderson MJ. *The American Census: A Social History*. New Haven, Conn: Yale University Press; 1988.
2. Evinger S. How shall we measure our nation's diversity? *Chance*. 1995;8:7-14.
3. Evinger S. How to record race. *Am Demographics*. 1996;18(5):36-41.
4. Edmonston B, Goldstein J, Lott J, eds. *Spotlight on Heterogeneity: The Federal Standards for Racial and Ethnic Classification. Summary of a Workshop*. Washington, DC: National Academy Press; 1996.
5. Standards for the classification of federal data on race and ethnicity. *Federal Register*. June 9, 1994:29831-29835.
6. Standards for the classification of federal data

on race and ethnicity. *Federal Register*. August 28, 1995:44674-44693.

7. Recommendations from the interagency committee for the review of racial and ethnic standards to the Office of Management and Budget concerning changes to the standards for the classification of federal data on race and ethnicity. *Federal Register*. July 9, 1997: 36874-36946.
8. Revisions to the standards for the classification of federal data on race and ethnicity. *Federal Register*. October 30, 1997:58781-58790.
9. *Review of Federal Measurements of Race and Ethnicity: Hearings Before the Subcommittee on Census, Statistics and Postal Personnel of the Committee on Post Office and Civil Service*, 103rd Cong, 1st Sess (1993).
10. *Testing Methods of Collecting Racial and Ethnic Information: Results of the Current Population Survey Supplement on Race and Ethnicity*. Washington, DC: US Dept of Labor, Bureau of Labor Statistics; 1996. BLS Statistical Notes 40.
11. *Findings on Questions on Race and Hispanic Origin Tested in the 1996 National Content Survey*. Washington, DC: US Dept of Commerce, Bureau of the Census; 1996. Population Division Working Paper 16.
12. *Results of the 1996 Race and Ethnic Targeted Test*. Washington, DC: US Dept of Commerce, Bureau of the Census; 1997. Population Division Working Paper 18.
13. *Racial and Ethnic Classifications Used by Public Schools*. Washington, DC: US Dept of Education, National Center for Education Statistics; 1996. Statistical Analysis Report NCES 96-092.
14. Office of Management and Budget. Draft provisional guidance on the implementation of the 1997 standards for the collection of federal data on race and ethnicity. Available at: <http://www.whitehouse.gov/OMB/inforeg/race>. Accessed February 17, 1999.
15. Office of Management and Budget. Guidance on aggregation and allocation of data on race for use in civil rights monitoring and enforcement. Available at: <http://www.whitehouse.gov/OMB/bulletins/b00-02.html>. Accessed March 9, 2000.