

Standards for the Classification of Federal Data on Race and Ethnicity



EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
WASHINGTON D.C. 20503

THE DIRECTOR

Federal Register, August 28, 1995

AGENCY: Executive Office of the President, Office of Management and Budget (OMB), Office of Information and Regulatory Affairs

ACTION: Interim Notice of Review and Possible Revision of OMB's Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting: Summary and Analysis of Public Comments and Brief Discussion of Research Agenda

Summary: In 1977, OMB issued the Race and Ethnic Standards for Federal Statistics and Administrative Reporting that are set forth in Statistical Policy Directive No. 15. The standards in this Directive have been used for almost two decades throughout the Federal government for recordkeeping, collection, and presentation of data on race and Hispanic origin. The standards have been used in two decennial censuses and in surveys of the population, data collections necessary for meeting statutory requirements associated with civil rights monitoring and enforcement, and in other administrative program reporting.

During the past several years, the standards have come under increasing criticism from those who believe that the minimum categories set forth in Directive No. 15 do not reflect the increasing diversity of our Nation's population. Some have also proposed changing the names of some categories. In response to the criticisms, OMB initiated a review of the Directive. As a first step in this process, OMB asked the Committee on National Statistics (CNSTAT) of the National Academy of Sciences to organize a workshop to discuss issues to be addressed in the review. A report of the workshop, held in February 1994, is forthcoming from CNSTAT. During 1994, the review process also included (1) public hearings in Boston, Denver, San Francisco, and Honolulu, (2) comment by Federal agencies on their requirements for racial and ethnic data, (3) development of a research agenda and related literature reviews, and (4) publication of a Federal Register notice, 59 Fed. Reg. 29831 (1994). The June 9, 1994, notice contained information on the development of the current standards and requested public comment on: (1) the adequacy of current racial and ethnic categories, (2) the principles that should govern any proposed revisions to the standards, and (3) specific suggestions for change that had been

offered by individuals and interested groups over the past several years. (See Appendix for the text of Directive No. 15.)

This Federal Register notice (1) summarizes the suggestions for changes drawn from public comments, research findings, and literature reviews, (2) briefly discusses the research agenda for some of the significant issues that have been identified, and (3) sets forth proposed principles to be used in reaching a final decision on standards for the classification of data on race and ethnicity. The issues, suggestions for change, and pros and cons described in this notice are those raised in public comment and do not reflect OMB positions or decisions. In addition it should be noted that because the categories in Directive No. 15 have been useful for over 18 years for many purposes, an option under consideration is to make no changes.

Important dates in the balance of the review process are shown below. Various agencies are conducting activities to support the review process; these include work by the Bureau of the Census related to the 2000 Census program mentioned below.

Fall 1995	OMB analyzes <i>Federal Register</i> notice comments; receives results of May 1995 CPS Supplement; continues to consult on options with affected groups
March 1996	Census Bureau conducts National Content Test (NCT) in preparation for 2000
June 1996	Census Bureau conducts Race and Ethnic Targeted Test (RAETT) in preparation for 2000 Census
November 1996	Bureau of the Census provides test through January 1997 results from National Content Test and Race and Ethnicity Targeted Test
Spring 1997	OMB publishes <i>Federal Register</i> notice on research results and proposed decisions on changes, if any, to Directive No. 15
Mid-1997	OMB publishes final decision regarding any changes to Directive No. 15 in a <i>Federal Register</i> notice

ISSUES FOR COMMENT: With this notice, OMB requests public comment on the following: (1) are there any issues or options not listed that should be considered before a final decision is made? (2) for each option presented, are there additional pros and cons to consider? (3) are there additional principles that should govern a final decision on whether or how to revise the standards? and (4) which options should be included for testing in 1996? This *Federal Register* notice provides the last opportunity for public comment on priorities for

All comments received as a result of the June 9, 1994, notice have been reviewed and considered in preparing this notice. It is not necessary to resubmit comments sent previously.

ADDRESS: Written comments on these issues may be addressed to Katherine K. Wallman, Chief, Statistical Policy, Office of Information and Regulatory Affairs, Office of Management and Budget, NEOB, Room 10201, 725 17th Street, N.W., Washington, D.C. 20503.

DATE: To ensure consideration, written comments must be provided to OMB on or before September 30, 1995.

ELECTRONIC AVAILABILITY AND COMMENTS: This document is also accessible on the U.S. Department of Commerce's FedWorld network under the "OMB Library of Files." The Telnet address for FedWorld via the Internet is fedworld.gov. The address (URL) for the World Wide Web is <http://www.fedworld.gov/ftp.htm#omb>. For ftp access, <ftp://fwux.fedworld.gov/pub/omb/omb.htm>. The telephone number for the FedWorld help desk is (703) 487-4608. For assistance in using electronic mail, please contact your system administrator.

Comments may be sent to OMB using the following Internet address: [ombdir15\(@\)a1.eop.gov](mailto:ombdir15(@)a1.eop.gov).

FOR FURTHER INFORMATION CONTACT: Suzann Evinger, Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, NEOB, Room 10201, 725 17th Street, N.W., Washington, D.C. 20503. Telephone: 202-395-3093.

SUPPLEMENTARY INFORMATION

A. Background

The United States government has long collected statistics on race and ethnicity. Such data have been used to study changes in the social, demographic, health, and economic characteristics of various groups in our population. Federal data collections, through censuses, surveys, and administrative records, have provided an historical record of the Nation's population diversity and its changing social attitudes and policy concerns. Since the 1960s, data on race and ethnicity have been used extensively in civil rights monitoring and enforcement covering areas such as employment, voting rights, housing and mortgage lending, health care services, and educational opportunities. These legislatively-based priorities created the need among Federal agencies for compatible, nonduplicative data for the specific population groups that historically had suffered discrimination and differential treatment on the basis of their race or ethnicity. In response, the Office of Management and Budget (OMB) issued in 1977 the "Race and Ethnic Standards for Federal Statistics and Administrative Reporting" contained in Statistical Policy Directive No. 15. These categories also implemented the requirements of Public Law 94-311 of June 16, 1976, which called for the collection, analysis, and publication of economic and social statistics on persons of Spanish origin or descent. Hence, the population groups identified by the Directive No. 15 racial and Hispanic origin categories reflected legislative and agency needs, and not efforts by population groups to be specifically identified.

In recent years, Directive No. 15 has been criticized for not sufficiently reflecting the Nation's diversity. In addition, some critics have proposed changing the names of some categories. In a June 9, 1994, *Federal Register* notice, OMB announced a review of Directive No. 15. As part of the review and public

comment period, OMB held hearings in Boston, Denver, San Francisco, and Honolulu. The June 9, 1994, *Federal Register* notice contains additional background information on the development of Directive No. 15; revisions proposed but not made in 1988; congressional hearings before the House Subcommittee on Census, Statistics, and Postal Personnel in 1993; a workshop conducted by the Committee on National Statistics in 1994; work done by the Interagency Committee for the Review of the Racial and Ethnic Standards;

In the June 9, 1994, *Federal Register* notice, OMB cited specific concerns the public had raised over the years regarding Directive No. 15. As a result of the notice, the public commented on the need for new categories, changes in current categories, whether racial and ethnic data should be collected, legislative and programmatic needs for the data, and the issue of self-identification versus observer identification. OMB received nearly 800 letters in response to the 1994 *Federal Register* notice and heard the testimony of 94 witnesses during the four public hearings. OMB heard from a wide array of interested parties including individuals, data users, and data providers from within and outside

This Federal Register notice focuses primarily on the six major issues discussed in comments from the public (Section B); the expected future research agenda (Section C); and general principles for making a final decision on standard racial and ethnic categories for Directive No. 15 (Section D).

Historical continuity of racial and ethnic data is important to many data users. Over time, however, there have been variations in how the Nation's principal population groups have been classified according to race and ethnicity; such differences have occurred even within data sets. In decennial censuses, for example, a question on race has been included since 1790. There have been many changes in the broad racial categories, the specific components of the categories, and whether data on ethnicity were collected. Asian Indians, for example, were counted as "Hindus" in censuses from 1920 to 1940, as "White" from 1950 to 1970, and as "Asians or Pacific Islanders" in 1980 and 1990.

Numerous studies reveal that identification of ethnicity is fluid and self-perceptions of race and ethnicity change over time and across circumstances for many people. This is especially true among persons with heterogeneous ancestries. A study of the Current Population Survey showed 1 in 3 people reported an ethnicity in 1972 that was different from the one they had reported in 1971. This level of inconsistency reflects the fluidity of ethnicity as well as the effect of question design.

Major historical inconsistencies in the data reflect social reality and public policy as well as technical decisions by data developers. Most agree that comparability over time is a desirable goal but that it is important also to reflect changes in society as they occur. Thus, General Principles 9 and 10 (see section D below) call for conducting research before any changes are made and for providing a crosswalk between old and any new categories so comparisons can be made across time.

There are also differences among data sets with respect to how race and ethnicity are classified. On birth records, for example, the race of the baby's mother and father are based on reports of the mother or family members. The race of the baby, which is not reported on the birth record, was once assigned for purposes of published statistics by an algorithm based on the parents' races. Since 1989, however, the National Center for Health Statistics has tabulated birth data according to the mother's race. In censuses and surveys until 1970, racial data were usually based on the observation of the government enumerator filling out the questionnaire. Now, the usual practice is self-administered forms and questionnaires, especially when the purpose of data gathering is to obtain information on population characteristics. In the enforcement of civil rights laws, however, the classification is often made by employers or school administrators, and the observer's perception is at issue. Whether someone is a victim of discrimination often turns on the way in which others act on their perception of, for example, the color of the individual's skin, the ethnic origin of his or her last name, or the accent with which he or she speaks. Such issues do not depend generally on the way in which the individual identifies his or her racial or ethnic background. In sum, Federal data sets identifying race and ethnicity are a mixture of self-identification by respondents and the perceptions of observers.

Until the current racial and ethnic standards were adopted in 1977, Federal data collections used an assortment of definitions for broad racial categories. In response to that problem, a Federal interagency committee recommended development of common categories for racial and ethnic data. Directive No. 15 provides a minimum set of standard categories and definitions for presenting data on various racial and ethnic groups in our population. The Directive requires compilation of data for four racial categories (White, Black, American Indian or Alaskan Native, and Asian or Pacific Islander), and an ethnic category to indicate Hispanic origin, or not of Hispanic origin.

To date evaluation of the quality of racial and ethnic data has been limited to research conducted by the Bureau of the Census, the National Center for Health Statistics (NCHS), and other parts of the Centers for Disease Control and Prevention (CDC). Comparisons of data sets indicate high consistency in individual responses for White and Black populations (95 percent consistency) and for the Asian and Pacific Islander population (90 percent consistency) in the 1990 census National Content Reinterview Survey conducted by the Census Bureau. For American Indians and Alaskan Natives, reporting is less consistent (63 percent consistency in the 1990 National Content Reinterview Survey). Reporting race is also less consistent for multiple-race persons, Hispanics, the foreign born, and persons who do not read or speak English well. NCHS found Asians and American Indians are sometimes misreported as "White" on death certificates, and this causes an underestimation of death rates for these groups. Nevertheless, these data quality problems are not so severe as to make the data unusable for most purposes.

Testimony at the four public hearings in 1994 and letters to OMB requested data on specific population groups that go beyond legislatively required levels of detail. Some groups say they have suffered

discrimination in political and economic access but without data for their specific population group, they feel that the discrimination is not recognized. For others, the request for recognition of a particular nationality group seems to be primarily a matter of pride and identification with that population group.

Public comment indicates self-identification is important to many people. Some who commented requested different placement of their specific group within a broad group. Many people of more than one race, who under Directive No. 15 are told to choose one category that "most closely reflects [their] recognition in [their] community," said they wanted to reflect their full heritage, not just part of it.

B. Summary of Issues and Suggestions Raised in Public Comment; Research Findings

In the June 9, 1994, *Federal Register* notice, OMB asked for public comment on (1) the adequacy of the current categories, (2) principles that should govern any proposed revisions to the standards, and (3) specific suggestions for changes that have been offered by various individuals and organizations.

This section summarizes the public comment (including comments from Federal agencies) that resulted from the June 9, 1994, *Federal Register* notice as well as research findings related to the particular issues. In an effort to be thorough in summarizing public comments the discussion below of specific data collection and presentation categories (Issue

The issues and suggestions shown below are those raised in public comment and do not reflect OMB positions or decisions. OMB will not make decisions on the issues until mid-1997. The following six issues are discussed in this section:

ISSUE 1.

Should the Federal government collect data on race and ethnicity? Should there be standards at all?

ISSUE 2.

Should Directive No. 15 be revised? Should there be different collection standards for different purposes?

ISSUE 3.

Should "race/ethnicity" be asked as a single identification or should "race" identification be separate from Hispanic origin or other ethnicities?

ISSUE 4.

Should self-identification or the perception of an observer guide the methods for collection of racial and ethnic data?

ISSUE 5.

Should population size and geographic distribution of groups be criteria in the final decision of Directive No. 15 categories?

ISSUE 6.

What should the specific data collection and presentation categories be? This discussion includes a brief summary of public comments and previous research findings. Briefly, suggestions that have been made include:

- (a) White (suggestions include adding categories for White ethnic groups; adding a category for persons from the Middle East or of Arab descent; and alternative wording for the category name).
- (b) Black (suggestions include identification of geographic origin of ancestors; adding a category for Creoles; and alternative wording for the category name).
- (c) Asian or Pacific Islander (suggestions include having three separate categories, one for Asians, one for Pacific Islanders, and one for Native Hawaiians; adding a new category for original peoples of acquired American lands ("indigenous populations") that would include American Indians, Alaskan Natives, Native Hawaiians, and native American Samoans and Guamanians; and specifying major nationality groups).
- (d) American Indian or Alaskan Native (suggestions include retaining the category with no change; expanding the definition of the category to include the Native Hawaiians and the indigenous populations of American Samoa and Guam; and alternative wording for the category name).
- (e) Multiracial (suggestions ranged from not having any multiracial category to six suggestions for ways to identify multiracial persons).
- (f) Hispanic origin (options include categories for subgroups; and alternative wording for the category name).

Detailed Discussion of the Six Issues

ISSUE 1. Should the Federal government collect data on race and ethnicity? Should there be standards at all?

Summary of views expressed on whether the Federal government should collect racial and ethnic data. Some agencies presently are required by Federal statute and regulation to collect racial and ethnic data. (See, for example, the Voting Rights Act of 1973 (1982) and the Civil Rights Act of 1964.) To end the collection of racial and ethnic data for these purposes, repeal of these statutes by Congress would

be required. The view of those who favor continued collection of racial and ethnic data can be summed up by the words of the writer who said, "...the measurable gains made in advancing a civil rights agenda to bring all Americans into the economic, political, and social mainstream would have been extremely difficult, if not impossible, if we did not have adequate information on racial and ethnic groups."

Those who favor no collection gave as their reasons the following: (1) doing so is divisive, archaic, unscientific, and racist; (2) it should not be a function of the Federal government (the government should be concerned only with citizenship) and the government has no need to know (tracking heritage is an individual choice and responsibility); (3) the government should collect ethnicity or ancestry instead of race; (4) there are no pure races, everyone is mixed, and therefore, the categories are meaningless; (5) people do not know their complete ancestry; (6) we are all supposed to have equal protection under the law (race neutral, color blind); (7) we are all Americans, we are a melting pot, we are one nation; (8) we are all human beings; (9) it is dehumanizing to categorize people like nuts and bolts; and (10) it is upsetting (for example, the categories are too limited; reminds people of the Nazi holocaust).

Should there be standards at all? Directive No. 15 is used widely and the strong consensus of public comment was to continue the issuance of standards for collecting data on race and ethnicity. The background and demand for the issuance of Directive No. 15 in 1977 is reviewed in 59 Fed. Reg. 29831, (1994).

As part of the public comment period, Federal agencies were asked to provide information about their requirements for data on race and ethnicity. Federal agencies report that the standards in Directive No. 15 have facilitated the exchange of data among agencies and among states, in instances where data are not used exclusively within a particular agency or program. Even where it is not required, Directive No. 15 standards are often used in State and business record systems and by marketers as a matter of convenience and to facilitate comparisons with other data sets.

The information also suggests, however, that Directive No. 15 may give a false sense of comparability and continuity among data sets. Even where the definitions of categories are comparable, there have been variations in collection and processing procedures that lead to inconsistencies in the data. Additional differences occur because of the mix of self-identification and observer-identification of race and ethnicity.

Agencies having statutory requirements to use racial and ethnic data for policy development, program evaluation, and civil rights monitoring and enforcement: (1) want historical continuity of the data; (2) generally oppose a "multiracial" category because the persons seeking this category are already covered by existing racial categories; (3) indicate that the perception of others is more valid for evaluating discrimination than individual self-identification; (4) note that standardized reporting formats,

like the Employer Information Report, EEO-1, rely on observer identification; (5) express concern about the cost of making changes that will affect both Federal agencies, respondents, and other governmental bodies; and (6) generally favor the broad group structure of Directive No. 15 in its present format.

Data collection agencies have legislative authority to collect racial and ethnic data needed for Federal programs and in the case of the decennial census, for redistricting. They also use racial and ethnic data for analyses of social, economic, and health trends for population groups. These agencies said: (1) the categories in Directive No. 15 confuse some respondents because they are inconsistent, too broad for some purposes, and the concepts of race, Hispanic origin, and ancestry overlap; (2) historical continuity of the data is important; (3) it is important to be able to aggregate any new categories back to the 1977 Directive No. 15 categories; (4) corrections are needed in Directive No. 15 (for example, there is no category for South American Indians and only Hispanic Whites and Hispanic Blacks are identified in the minimum combined format); (5) subgroups of Asians and Hispanics were most frequently cited as a need but required data collection should be limited to groups with sufficient numbers to generate meaningful estimates; (6) a few agencies expressed interest in subcategories of the Black population (e.g., African, West Indian); and (7) for American Indians, some expressed a need to require the identification of Federal- versus state-recognized tribes. Many felt a "multiracial" category (that does not specify the races) is too heterogeneous and affects the counts of other groups in unknown ways. Agencies that collect health data particularly need to know specific categories because some diseases and health problems are more prevalent among certain racial and ethnic groups. Data collection agencies are concerned about the significant operational, technical, and cost issues of a "check all that apply" approach for multiracial persons. For example, processing systems would have to be changed to allow for reporting more than one category. Additionally, Federal laws have been written with the assumption that persons identify with one racial group; these laws would either have to be changed or some method would have to be devised to meet legislative requirements.

Federal agencies have interpreted Directive No. 15 to apply only to primary data collection; data collection under grants may or may not comply with it.

ISSUE 2. Should Directive No. 15 be revised? Should there be different collection standards for different purposes?

Among those who favor collection of racial and ethnic data, there is significant difference of opinion as to whether Directive No. 15 should remain essentially as it is or should be revised. While some believe there should be no change in Directive No. 15, others say ethnic identification is in constant flux and Directive No. 15 should be changed now and subsequently reviewed periodically (for example, after every decennial census). The Directive No. 15 categories are nearly two decades old and many people say they no longer identify with the categories. Intermarriage, changes in immigration flows, and changes in ethnic consciousness are some of the reasons. These changes in our basic population

structure suggest an increasingly diverse society and unforeseen future needs for racial and ethnic data.

Public testimony and research indicate that race and ethnicity are subjective concepts and inherently ambiguous. For purposes of collecting data in the United States, race and ethnicity are cultural concepts and social constructs. As stated in the current version of Directive No. 15, the racial and ethnic categories are not intended to reflect scientific or anthropological definitions of who should be included in a particular category. The definitions of the minimum set of population categories under Directive No. 15 include references to color, ancestry, and geographic origins in an effort to approximate social constructs of race prevalent in the United States.

In line with the subjective nature of the concept, research shows people change how they classify themselves with respect to race and ethnicity. There is significant inconsistency in the measurement of ethnicity particularly. Research shows different responses are summoned by the format of questions (open or specified categories), the number of categories, the examples listed, changes in self-perceptions within groups and among age cohorts, and the political climate.

The differing views of whether Directive No. 15 should be revised relate to the purpose for collecting such data. Federal agencies that use racial and ethnic data for regulatory programs, civil rights monitoring and enforcement generally oppose any revision of Directive No. 15 for the reasons described in Issue 1. Directive No. 15 is seen as providing practical guidelines for visual identification in a broad and relatively straightforward manner of the population groups that have historically suffered discrimination.

Where trend analysis of social and economic changes was the commenter's purpose, more detailed categories were often favored. The preference varies for other purposes such as policy development and program fund allocations. In the public hearings and letters to OMB, persons concerned with self-identification generally favored revisions that would provide more detailed categories and more freedom of choice (see Issue 6).

Given the distinct uses of racial and ethnic data in the Federal government (especially trend analysis versus regulatory and civil rights monitoring and enforcement), the possibility of a two-part Directive No. 15, with one part focusing on each purpose, has been suggested as an option if there are changes to Directive No. 15. Part A of Directive No. 15 could provide more detailed standards for use when a major purpose is trend analysis (such as in the decennial census and perhaps household surveys). Such a standard would track the increasing diversity of the U.S. population and provide better information to inform decisions about whether the categories for administrative and enforcement purposes should be expanded. Part B of Directive No. 15 could remain essentially unchanged for use in program evaluations and civil rights monitoring and enforcement.

There are disadvantages to having two levels of data collection specified in the standards of a revised Directive No. 15. The most serious disadvantage could be data sets with different counts of population groups that cannot be related, a result of different coding and tabulation rules. This is especially the case if the specific races of multiracial persons are identified. Two sets of data could be confusing to data users who may be unsure of which set to use for various purposes. To prevent refocusing the problem from data collection to tabulation, there would have to be generally agreed-upon procedures and guidelines for how agencies would tabulate data for program purposes. The procedures should ensure that detailed data collections could be tabulated back to the broad categories of the 1977 Directive No. 15 in a standard way across programs. Standard and generally agreed-upon tabulation rules would be needed for the various combinations of multiracial entries, including those where neither race is "White." The Bureau of the Census already has procedures for aggregating detailed data from the 1990 census to the broader categories of Directive No. 15. The reaggregations could become more complicated because of the different assumptions that would be required. The requests of some groups who do not feel they fit into existing categories (e.g., some Arabs, Creoles, and Cape Verdeans) suggest that aggregations could become even more problematic. Also, the quality of the reaggregated data can vary by geographic area.

Some say cost should not be an "excuse" for failing to improve data collection on race and ethnicity, especially where the data are used for protection of civil rights. Others expressed concern about the cost of making changes to Directive No. 15 when the broad categories are acceptable choices for most of the population and cover programs affecting almost all persons. Added costs associated with more detailed categories are discussed in Issue 6 below.

Federal, State, and local government agencies urged that any revisions ensure that data can be tabulated back to the 1977 categories. Most expressed a preference to maintain historical continuity of the two decades of data sets with the understanding they are not perfectly comparable. It was also recognized that final tabulations give the data an appearance of comparability among data sets when actually there are differences caused by data collection methods (especially self-identification versus identification by observers). Nevertheless, the data are widely accepted by courts and government agencies as reliable indicators of change in housing patterns, redistricting, and labor markets.

If there are revisions to Directive No. 15, research indicates that changes in the race and ethnic categories on administrative records will present problems in data comparability over time. The categories on the records reflect what they were as of the time of initial enrollment and the categories are generally carried without change for decades. Administrative records are often collected from State and local sources, which have a variety of recordkeeping practices, are not required to meet Directive No. 15 (but often do), and are unlikely to collect information for detailed categories. A few States now require a "mixed race" category. There will be increasing value to the Federal government if State records use the same categories as Directive No. 15.

Federal and State government agencies emphasized that if there are revisions, a reasonable amount of time needs to be given to phase in the changes.

ISSUE 3. Should "race/ethnicity" be asked as a single identification or should "race" identification be separate from Hispanic origin or other ethnicities?

Directive No. 15 states that it is preferable to collect data on race and Hispanic separately to allow flexibility. If a combined format is used to collect racial and ethnic data the minimum acceptable categories are: American Indian or Alaskan Native; Asian or Pacific Islander; Hispanic; White, not of Hispanic origin; and Black, not of Hispanic origin. The use of the Hispanic category in the combined format does not provide information on the race of those selecting it. As a result, the combined format makes it impossible to distribute persons of Hispanic ethnicity by race and, therefore, reduces the utility of the four racial categories by excluding from them persons who would otherwise be included. Thus, the two formats currently permitted by Directive No. 15 for collecting racial and ethnic data do not provide comparable data.

Public testimony reflected some data problems with the standards in Directive No. 15. The combined format does not provide for identification of Asians or American Indians with Hispanic origins, and would classify the people of Equatorial Guinea, who are geographically Africans but who speak Spanish, as Hispanic. There is no apparent category for Central and South American Indians.

Some persons from non-Hispanic ethnic groups questioned why Hispanics had been singled out as the only ethnic group specifically identified in Directive No. 15. Others objected to the term "non-Hispanic" because it defines people by what they are not. For example, rather than "White, not of Hispanic origin," a category might be "White, European ethnicity" or "American Indian, Mexican." This approach would require a question that identifies ancestry groups within the broad race groups.

Most Federal agencies did not comment on whether race and Hispanic origin should be collected in one question or two questions, although many agencies have been using the combined format for a number of years and have developed data series with the resulting data. Those few that commented were split on the issue.

The public indicated differences of opinion also. Those who favored asking race and Hispanic origin separately said Hispanics were a multiracial population and a cultural (not a race) group. Many Latin American countries are populated by immigrants from parts of Europe other than Spain. Many wanted to identify Asian-Hispanics and American Indian-Hispanics. Research shows Hispanics who self-identify as White also fare better economically; thus, some said two questions were needed because ethnicity alone was insufficient for determining which Hispanics are likely to be victims of discrimination. Others were concerned with historical continuity of data concepts and wanted to be able to generate

statistics for the total White and total Black population. When separate questions are used to collect racial and ethnic data, there is also a technical matter of which question should be asked first.

Some who favored asking race/Hispanic origin as one question said many Hispanics do not identify themselves as a race. Others favored this approach as a way to end the practice of using the term "race" which they see as a social rather than a scientific construct.

For some individuals, race and ethnicity may not be clearly separable. One proposed solution is to ask a single race/ethnicity question (that is, one question in which "Hispanic" is included in the list with the broad race categories) and allow respondents to mark all that apply. Hispanics who identify with a race category could mark both categories. Hispanic respondents who do not identify with any race category could mark "Hispanic" only. The question would correspond to self-perceived membership in population groups defined by cultural heritage, language, physical appearance, or other characteristics.

Some research supports the public comments that some respondents are confused about how to respond to separate race and Hispanic origin items. In the 1990 census, 4 in 10 Hispanics marked "Other" in the race question and about 10 percent of the population did not respond to the Hispanic origin item. The 1990 census reinterview study, in which the answers given by a sample of respondents to the 1990 census were compared with answers they gave in a reinterview after the census, also showed that Hispanics had high levels of inconsistent reporting in the race item. These results indicate the question may not be operating as intended.

Cognitive research shows that many Hispanics perceive redundancy in separate race, Hispanic origin, and national origin questions. Some Hispanic respondents do not identify with the Black or the White category, and are offended by an "Other race" category (which they interpret to mean that Hispanics are less important than other races since they do not have their own "label"). For some, "White" is synonymous with "Anglo" meaning non-Hispanic. For example, in a focus group, a Mexican-American man said that where he lived people were either Mexicans or Anglos. He was confused by a race question that seemed to be trying to make him say he was White and to his mind, non-Hispanic. In an analysis of the responses of Hispanics to the race question in the 1990 Panel Study of Income Dynamics, Cubans were the most likely and Mexican-Americans the least likely to identify themselves as "White." Cognitive research shows some Hispanics, especially the foreign born, expect to see a single category for Hispanics.

If race and Hispanic origin are asked as two separate questions, there is the issue of whether to ask race or Hispanic origin first. Research done since 1987 indicates that additional instructions and asking Hispanic origin first reduce nonresponse to that question. Asking Hispanic origin first also reduces reporting as "other race" and increases reporting as "White" by U.S.-born Hispanics but not by immigrants. A large minority of respondents still report as "other race." The Census Bureau will conduct research in the 1996 National Content Test for the 2000 census to determine whether placing the

Hispanic item first affects consistency of responses and reporting in the race category among subgroups not adequately represented in other studies.

The future research agenda is described in Section C below.

ISSUE 4. Should self-identification or the perception of an observer guide the methods for collection of racial and ethnic data?

At the heart of criticisms and public requests for review of Directive No. 15 is the feeling of some persons, particularly those of mixed heritage, that they cannot accurately identify their race and ethnicity as they prefer in Federal data systems using the current categories. They say the government should not limit their choice of identification. As stated in the second principle for the review of racial and ethnic categories (Section D below), ideally OMB prefers that self-identification should be facilitated to the greatest extent possible but there are data collection systems where observer identification is more practical. Federal censuses, surveys, and vital records give preference to using self-identification; that is, having the individual (or in some cases a proxy respondent) provide the information requested about his or her race and Hispanic origin.

Research shows that ethnic groups evolve and may modify their preferred ethnic group names; individuals may represent their affiliation with groups differently depending on the situation and may alter their perceived ethnic membership over time. Category names need to be acceptable and generally understood both by members and nonmembers of the groups to which they apply.

Self-identification is not the preferred method among Federal agencies concerned with monitoring and enforcement of civil rights. They prefer to collect racial and ethnic data by visual observation. Since discrimination is based on the perception of an individual's race or Hispanic origin, these agencies oppose any changes that would make it more difficult to collect data by observation. Such proposed changes include the suggested "multiracial" category as well as identification of national origins and ethnicities (for example, "Arab" or "Cape Verdean"). These agencies say that if categories are more detailed and include nationality groups, or if there is a "multiracial" category (and especially if the multiple races have to be identified), it would be virtually impossible to give instructions for how to classify by visual observation. Additionally, they report it is their experience that direct inquiry about a person's race, ethnicity, or national origin sometimes raises concerns among employees or other respondents about the purpose of collecting the data.

American Indian groups express concern about self-identification. Tribal recognition of status as an American Indian or Alaskan Native (Alaskan Indian, Eskimo, or Aleut) is a legal definition, not one of long-ago ancestry. In the 1990 census, 8.7 million persons reported in the ancestry question that they were American Indian but only 1.9 million reported American Indian race. Only 3 of 4 who reported "American Indian" as their race gave "American Indian" as their first ancestry; about 9 percent gave an

European first ancestry. There are also regional effects in reporting American Indian as a race related to the prevalence of intermarriage, migration, Federal recognition of regional tribes, and attitudes towards Indians.

Development of Federal data sets includes increased use of administrative records matched to survey data for trend analysis. This makes the issue of data collection methods, both by observation and self-identification, a greater technical difficulty than in the past. Where identification is by observers or proxy respondents, blood relatives may be identified differently in administrative records and an individual may be identified differently among data sets.

ISSUE 5. Should population size and geographic distribution of groups be criteria in the final decision of Directive No. 15 categories?

Many of the groups for which data collection has been requested are numerically small and often are found primarily in specific geographic areas. In national sample surveys, these factors often make it unreasonably costly or burdensome on the public to collect reliable data. A question that allows for self-identification to the greatest extent possible may be very lengthy. Some see this as a technical problem, others do not.

There are difficulties with using size of population as a basis for making a population group a specific category. The size of the population is itself a subject of controversy at times.

For sample surveys, how small is "too small"? Sample data can provide only an estimate of a number and not, with 100-percent certainty, the true number itself. The smaller the group, the more unreliable estimates are with respect to sampling error. For example, in the Current Population Survey (CPS), a national survey of households, summary measures such as means and percentage distributions are shown only when the population base is 75,000 or greater. An example of how much sampling error increases in a survey as the population size of a group decreases can be provided for a characteristic such as the poverty rate. If the estimated poverty rate for the total U.S. population is about 14 to 15 percent (a 90-percent confidence interval), then for a population group of 1 million persons, the poverty rate would be about 8 to 21 percent; for a population group of 500,000 persons, the poverty rate would be about 6 to 23 percent; and for a population group of 200,000 persons, the poverty rate would be about 1 to 28 percent. (A 90-percent confidence interval can be interpreted roughly as providing 90-percent confidence that the true number falls between the upper and lower limits.) The accuracy and reliability of an estimate depends not only upon sample sizes, but also upon whether the groups are "controlled" (i.e., weighted to independent estimates). Estimates of the Asian and Pacific Islander population from the 1994 March Current Population Survey differed by about 20 percent from demographic estimates due primarily to this factor.

One person suggested that groups should constitute at least one percent of the population (nationally, about 2.6 million in 1994) to be considered as a separate category. A time frame and data source would have to be agreed upon if such a guideline were considered.

ISSUE 6. What should the specific data collection and presentation categories be?

There are no clear, unambiguous, objective, generally agreed-upon definitions of the terms, "race" and "ethnicity." Cognitive research shows that respondents are not always clear on the differences between race and ethnicity. There are differences in terminology, group boundaries, attributes, and dimensions of race and ethnicity. Historically, ethnic communities have absorbed other groups through conquest, the expansion of national boundaries, and acculturation.

Groups differ in their preferred identification. Concepts also change over time. Research indicates some respondents are referring to the national or geographic origin of their ancestors, while others are referring to the culture, religion, racial or physical characteristics, language, or related attributes with which they identify. The 1977 Directive No. 15 categories are a mix of these. The categories do not represent objective "truth" but rather, are ambiguous social constructs and involve subjective and attitudinal issues.

Some said the categories should reflect ancestry or cultural affiliation rather than skin color. Some wanted to indicate they were "American" and had ancestry from a particular geographic region ("hyphenated Americans") while others opposed this ("we are all Americans"). Cognitive research indicated that some people use race and ethnic origin interchangeably; they see little difference between the two concepts. Most people do understand the concept of ancestry.

Some groups stated that their preference was for standard categories that would maximize the size of their population because they believed larger numbers provide importance in society and greater political leverage.

In short, groups differed in what they considered the most desirable standard. It is impossible to satisfy every request for racial and ethnic categories that OMB received; such a list would be both lengthy and contradictory. Some persons requested religious identification; this option is not discussed below because the Federal collection of religious affiliation has been interpreted as possibly violating the separation of church and state.

Some suggested a completely open-ended question with no standard categories for data collection; rather, standards would be set for data tabulation. An open-ended question is discussed in part (e), Multiracial option (2)(cc).

Below is a discussion of public comment with regard to the current broad categories of "White," "Black," "Asian or Pacific Islander," "American Indian or Alaskan Native," and "Hispanic origin." Part (e) below

discusses options with respect to classification of persons of multiple races, a category that does not exist in the current standards. Where possible, in the discussion of options and their pros and cons, past research results are included.

As part of the discussion of options, the cost of proposed changes with respect to collecting, tabulating, and analyzing data is an essential consideration (see Section D, General Principle 8). Any changes in Directive No. 15 will be imposed on tens of thousands of State and local agencies such as law enforcement agencies (through the Uniform Crime Reporting system), school districts, the business community, and others required to use the Directive in reporting these data to the Federal government. If administrative records for Federal programs have to be completely updated to meet a new standard, there will be significant costs to entities that report to the Federal Government. For example, the State of Florida estimates it would cost \$2 million to change school enrollment records.

Changes in the current Directive No. 15 would also entail additional processing costs as software and sometimes data capture methods would have to be changed. For example, it is more expensive to capture and code handwritten responses to open-ended questions than fixed, pre-determined categories. Some of the increased costs associated with categories more detailed than the current Directive No. 15 would include:

- interviewer training for implementing changes in collecting these data and updating of interviewer instruction manuals;
- additional interview time to collect more detailed data;
- the technical and practical difficulty of administering more detailed or more complex categories (such as long lists of nationalities, especially if multiple responses are allowed) in telephone surveys;
- increases in computer reprogramming and data processing costs;
- increases in the likelihood of litigation over data aggregation and processing decision rules;
- increases in the costs of disseminating data in hard copy or electronic format and storing larger computer data files;
- updating program manuals, regulations, and recordkeeping requirements to reflect changes; and
- making data analysis more complex.

The cost considerations described above apply, in varying degrees, to any change and so are not described further in the discussion below of pros and

(a) White

In Directive No. 15, the "White" category includes persons having origins in any of the original peoples of Europe, North Africa, or the Middle East. The public comment included suggestions for subcategories and related changes in terminology to collect more detailed information on White ethnic groups according to the geographic region of their ancestors. This summary reports only on options proposed during public hearings and in the public comment period. It also highlights pros and cons for these options as raised in public comment or shown by research. Inclusion in the summary does not reflect OMB endorsement of the comments or suggestions. Requests included:

Options Suggested in Public Comments:

(1) Collect data for White ethnic groups according to the country of ancestral origin (for example, German, Scottish, or Irish). Some prefer other terms such as "European-American," or "German-American" and some requested that "European" be further subcategorized into "Western European" and "Eastern European." Some suggested subcategories for identifying the original peoples of Europe, North Africa, and Southwest Asia (Middle East).

Pros of Option (a)(1):

- Collection: Some persons identify more with their ancestry than with "White" as a racial category.
- Tabulation and analysis: Indicates the ethnic diversity of the "White" population.

Cons of Option (a) (1):

- Collection:
 - Physical space on forms: if national origin groups are listed, considerably more space would be needed.
 - Telephone interviews: more difficult than currently, especially if national origin groups are listed; more time consuming to ensure that respondent is given the opportunity to understand the possible choices.
 - Data quality: effect on counts of specific nationality groups if country not listed as an example; count also affected by which nationalities/countries are listed; respondent may be confused between reporting ancestry (e.g., German) versus country of birth (e.g., Russia); and some respondents do not know their

ancestry.

--Categories not required by Federal legislation/programs and States unlikely to collect this detail in administrative records.

--Visual observation: nationality not easily determined.

- Tabulation and analysis: Need rules for tabulating multiple ancestries. More categories add costs for tabulation and analysis. It is more complex to analyze and report on many nationalities as compared with single race categories.

(2) Create a separate category for Arabs/Middle Easterners (currently included as part of the "White" category) in order to distinguish this population from persons of European descent in the "White" category. The public comment offered different suggestions for the name of the category and how to define the population group it would be intended to cover. Some comments supported a separate category for the decennial census enumeration, but not necessarily adding a separate category to the minimum set of racial and ethnic categories in Directive No. 15. These suggestions included:

-- Create a geographically oriented category called "Middle Eastern" (based not on race but on region of origin) for persons from the Middle East/North Africa and West Asian region, regardless of their race, religion, or language group. It would include Arab states, Israel, Turkey, Afghanistan, and Iran. Some suggested also including Pakistanis and Asian Indians in their geographic definition of the term. Data availability on subsets of the Middle Eastern regional category was also requested. Some comments referred to the "Middle Eastern" category as an ethnic identifier; some favored the addition of a "Middle Eastern" category to the list of basic racial and ethnic categories; and others suggested a "Middle Eastern" subcategory be created within the "White" category. Those preferring a "Middle Eastern" to an "Arab" category felt that the category would build on the other regionally defined categories, consolidate people from different countries but with similar cultural/geographic experiences regardless of race, and distinguish them from persons of European descent in the "White" category.

-- Add an ethnic category called "Arab-American" based on a linguistic and cultural approach to the minimum set of categories in Directive No. 15. Those who preferred the term, "Arab" said Arabs, like Hispanics, are an ethnic group of mixed race and have a shared language and culture. They would make "Arab" a separate category rather than part of the "White" category; they would leave North Africans, who are not Arabs, as part of the "White" category.

-- Reclassify "Muslim West Asians" as part of the "Asian or Pacific Islander" category.

Pros of Option (a)(2):

- Collection:

--A separate category would satisfy Arab/Middle Eastern respondents who do not think of themselves as "White" or as having any identity in common with Europeans.

--A separate category would facilitate self-identification and could possibly improve the quality of the data on Arabs/Middle Easterners.

--Telephone survey: Easy to ask if it is the only category added; however, if additional categories are added, it may be problematic.

- Tabulation and analysis:

--Would provide treatment comparable to Hispanics (and in some data sets, specific Asian nationality groups).

--Data could be used in policy development, in delivery of services and needs assessments, for civil rights monitoring and enforcement, and in health research.

--Reflects the ethnic diversity of the "White" category.

Cons of Option (a)(2):

- Collection:

--Requires space on form for an additional category.

--There was no general agreement in public comment about the geographic definition of "Middle East." For example, there is disagreement in public comment about whether Pakistanis and Asian Indians are included if the term, "Middle Eastern" is used. The term, "Arab" clarifies that Asian Indians and Pakistanis would remain classified with Asians, which some consider preferable for historical continuity; no requests were received from Asian Indians or Pakistanis to be reclassified. Public testimony indicated inconsistencies in understanding which countries should be included as "Arab."

--Identification by observers: because some Arabs are light-skinned and some are dark-skinned,

identification by visual observation is prone to error.

- **Tabulation and analysis:**

--A separate Arab/Middle Eastern category may affect the historical comparability of data in the "White" category and may affect the counts of other racial groups since Arabs are a mixed racial group.

--Adds a category on many national surveys for a geographically concentrated population (about half of the Arab population is concentrated in Detroit, New York, Los Angeles, and Washington, D.C.).

--Relatively small population (less than 900,000 according to the 1990 census, although Arab groups claim 1-3 million). Geographically, "Middle Easterner" as a category would include persons other than Arabs.

(3) Alternative words suggested for "White" include "Caucasian" and "Anglo."

Pros of Option (a)(3):

- Collection: "Caucasian" preferred by some respondents.

Cons of Option (a)(3):

- Collection: "Anglo" generally refers to Whites of European ancestry and excludes Hispanics; this would affect historical continuity. The term tends to be used regionally and may not be generally understood.

Past research results/literature review: Some object to the term "White" (for example, in cognitive research one said, "white is the color of paint" and in a letter another said, "I am not the color of this paper"). Some preferred the term, "Caucasian." Ethnicity is largely symbolic or optional for many Whites. Whites often reported inconsistently, as "American," or not at all in response to the 1990 census ancestry question. A significant number of Whites do not strongly identify with a specific European ethnicity. This has been the case for decades. For example, only about 55 percent of matched persons who reported English, Scottish, or Welsh in the March 1971 Current Population Survey (CPS) reported the same origin in March 1972. The "example effect" is very strong for White ancestry groups. For example, in two surveys held five months apart, 40 million people reported English as their ancestry and in the other, nearly 50 million said they were English. The only difference was placement of a question on language use in their home (English for 90 percent of the population) after the ancestry question in the second survey and farther apart in the first survey. "German" was the first example in the 1990 census ancestry question and, as a result, the German population appeared

to grow very rapidly. Some Whites, however, do identify strongly with their ancestry and were confused by the 1990 census race question which listed nationality groups for Asians and Hispanics but not for Whites.

(b) Black

The term "Black" in Directive No. 15 refers to a person having origins in any of the Black racial groups of Africa. There were suggestions to change the definition to "persons having origins in any of the Black peoples of Africa," or to define the term to include all Black persons regardless of country of origin or country of citizenship. Requests were made to identify Blacks according to the geographic region of their ancestors. "African-American" and "Black African-American" were suggested as names for the category [the suggestions of "Black American" and "Amerofian" (described as Blacks who are American Natives, European, and West African) are not discussed below]. This summary reports only on options proposed during public hearings and in the public comment period. It also highlights pros and cons for these options as raised in public comment or as shown by research. Inclusion in the summary does not reflect OMB endorsement of the comments or suggestions. Requests included:

Options Suggested in Public Comments:

(1) Collect data for Black ethnic groups according to geographic origin of Black ancestors (African, Haitian, Jamaican, Caribbean, West Indian, Brazilian, Ethiopian, etc.).

Pros of Option (b)(1):

- Collection: Easy to ask in a telephone survey. Some persons identify more with their ancestry than with "Black" as a racial category.
- Tabulation and analysis: useful for research on health, diversity, needs assessments, trends analysis; does not affect historical continuity. Indicates the ethnic diversity of the Black population

Cons of Option (b)(1):

- Collection:

--Determining geographic origin or nationality/ancestry by visual observation would be difficult and prone to error.

--If national origin groups are listed, considerably more space would be needed on forms.

--Telephone interviews would be more difficult than currently, especially if national origin groups are listed; more time consuming to ensure that respondent is given the opportunity to understand the possible choices.

- Tabulation and analysis: Data not needed for Federal program evaluation and enforcement. States are unlikely to collect this detail in administrative records. Count of specific nationality groups could be affected if respondent is confused between reporting ancestry (e.g., Haitian) versus country of birth (e.g., Virgin Islands); some respondents do not know their ancestry. Rules would be needed for tabulating multiple ancestries. More categories add costs for tabulation and analysis. It is more complex to analyze and report on many nationalities as compared with single race categories.

(2) Create a separate category for Louisiana (French) Creoles. They objected to categorization with Blacks as they are a multiracial/ethnic group (African, French, American Indian, and Hispanic).

Pros of Option (b)(2):

- Collection: Easy to ask in a telephone survey if it is the only category added; however, if additional categories are added may be problematic.

Cons of Option (b)(2):

- Collection: Extra space on forms; extra time in telephone interview.
- Tabulation and analysis: Small population size. Affects historical continuity of data sets.

(3) Use the alternative term, "African American" or "Black, African-American."

Pros of Option (b)(3):

- Collection: Commonly-used identification for Blacks born in the United States or Blacks whose parents are Americans.

Cons of Option (b)(3):

- Collection:

--Excludes Africans who are not American citizens.

--Term generally refers to Blacks but respondents could reasonably be confused as to whether to include Whites, Asians, or others born in Africa (especially North Africa). Could affect historical comparability of data.

--Blacks born in Brazil or the Caribbean (especially immigrants) do not identify with the term, "African American." Some Blacks who have been in the United States for generations have no record of where in Africa their ancestors were born and do not wish to be called "African-Americans."

--Use of "American" increases respondent error by persons who are not Black but who wish to identify as "Americans."

- Tabulation and analysis: refers to a continent, not a country.

(4) Provide a separate category for Cape Verdeans (Portuguese and African ancestry from Cape Verde on the western tip of Africa. This is mostly a multiracial population. "Cape Verdean" is generally considered a national, ethnic and linguistic designation rather than a racial designation). The category could be an ethnic category rather than a racial category as is the case for persons of Hispanic origin.

Pros of Option (b)(4):

- Collection: Would satisfy Cape Verdean respondents and is easy to ask.
- Tabulation and analysis: useful for civil rights monitoring and enforcement in State of Massachusetts.

Cons of Option (b)(4):

- Collection:

--Visual identification of Cape Verdeans prone to error because of various skin colors. People within the same family say they are identified differently.

--Adds a category for a small, geographically concentrated population.

- Tabulation and analysis: not required for Federal programs.

Past research results/literature review: In surveys from 1989 to 1991, more Blacks said it did not matter if they were called "Black" or "African American" than said they preferred one over the other. Among those with a preference, the ratio choosing "Black" over "African-American" was 1.2 to 1. In a 1993

survey in the Chicago area, a majority of Blacks preferred "African American" for their ethnicity and "Black" for their race but the proportion had declined since 1991.

Several studies of Blacks with roots in the Caribbean or Africa show they do not feel they share a common history or culture with American-born Blacks and distinguish themselves from this population. Further research is needed on the terminology that is generally understood or most acceptable. In the 1990 census, about 370,000 persons wrote in an entry classified as "Black"; about three-fourths of these were ethnic subgroups such as Jamaican and Haitian. Cognitive research suggests that many foreign-born Blacks interpreted the race question in terms of national origin rather than race.

(c) Asian or Pacific Islander

The definition used for "Asian or Pacific Islander" in Directive No. 15 refers to a person having origins in any of the original peoples of the Far East, Southeast Asia, the Indian subcontinent, and the Pacific Islands. This area includes, for example, China, India, Japan, Korea, the Philippine Islands, the Hawaiian Islands, and Samoa. Public comment indicated confusion about which countries are included in this definition, particularly for "Indian subcontinent" and whether the aboriginal peoples of Australia are included in this category. Requests were made to have separate categories for Asians and Pacific Islanders and to provide additional subcategories under "Asians" to describe better this diverse population; to move Native Hawaiians, American Samoans, and Chamorros to either a separate category or to the "American Indian or Alaskan Native" category. This summary reports only on options proposed during public hearings and in the public comment period. It also highlights pros and cons for these options as raised in public comment or shown by research. Inclusion in the summary does not reflect OMB endorsement of the comments or suggestions. Requests included:

Options Suggested in Public Comments:

(1) Make two categories, one for "Asians" and one for "Pacific Islanders." Pacific Islanders include indigenous populations from American Samoans, Carolinians and Chamorros, and Native Hawaiians, as well as other population groups in the Pacific Islands. Native Hawaiians have a specific legal status in Federal statutes different from other indigenous Pacific Islanders.

Pros of Option (c)(1):

- **Collection:**

--Easy to ask in a telephone survey.

--Categories are mutually exclusive.

- Tabulation and analysis:

--Pacific Islanders are culturally and ethnically distinct from Asians so separate data would be useful for trends analyses, needs assessments, and health research. Historical continuity can be maintained by aggregating "Pacific Islanders" with "Asians."

--Separate categories for Pacific Islanders and for Native Hawaiians would meet program needs of the Department of Veterans Affairs to report on veterans from specific minority groups.

Cons of Option (c)(1):

- Collection:

--Adds a category.

--Respondents may be confused as to the exclusion or inclusion of Native Hawaiians and other indigenous populations which could seriously affect data quality.

--Effect on data collected by visual observation is unknown.

- Tabulation and analysis:

--Pacific Islanders are geographically concentrated and a relatively small population group for a separate category.

(2) Specify major Asian nationality groups.

Pros

- Collection: Done successfully in the 1990 census.

- Tabulation and analysis: Indicates diverse and significant differences in the characteristics of the Asian population; potentially useful in analyses of health and other trends.

Cons of Option (c)(2):

- Collection:

--Requires significant physical space on forms.

--Telephone interviews: tedious to read long lists.

--Identification by observers: difficult to determine specific nationality.

(3) Develop a new category for original peoples of acquired American lands ("indigenous" populations). This would include persons having origins in any of the original peoples of North America who maintain cultural identification through tribal affiliation or community recognition (American Indians, Alaskan Indians, Aleuts, and Eskimos); the Hawaiian Islands; American Samoa; Guam; and the Northern Marianas. Some suggested this be a "Native American" category. Refer also to Option (d)(2) below.

Pros of Option (c)(3):

- Collection:

--Many Native Hawaiians preferred this option. They do not consider themselves Asians and they insist that they are not immigrants to the United States. They said that including them in the large "Asian or Pacific Islander" category resulted in data that do not accurately reflect their social and economic conditions. Some representatives of Asian groups supported this suggestion.

--No increase in the number of categories.

--Category mutually exclusive.

- Tabulation and analysis:

--Inclusion of indigenous Pacific Islanders as "Asians or Pacific Islanders" masks their economic status. For example, Pacific Islanders have relatively high poverty rates. They also have health issues and educational needs different from Asians.

Cons of Option (c)(3):

- Collection:

--This might be viewed as a political category rather than as one commonly recognized by most individuals in society.

--Respondent error likely both on forms and in telephone surveys as "indigenous" or "original peoples" are not familiar terms to most of the population; the term, "native" is interpreted to mean any person born in a particular area. No generally-understood choice for the category name.

--Unknown how data collected by visual observation would be affected.

- Tabulation and analysis:

--Opposed by most American Indian tribal governments and organizations as they preferred to maintain a category which refers specifically to American Indians.

--Heterogeneous population in terms of characteristics; data would be less useful than currently for policy development, trend analyses, and needs assessment. Not useful for health research. American Indians were particularly concerned about possible effects on the quality of data needed for programs and funding.

(4) Have a separate category for Native Hawaiians (defined as individuals who are descendants of the aboriginal people who, prior to 1778, occupied and exercised sovereignty in the area that now constitutes the State of Hawaii). Change "Hawaiian" to "Hawaiian, part-Hawaiian," because most Native Hawaiians are part Hawaiian and many, in the past, have categorized themselves as "White."

Pros of Option (c)(4):

- Collection:

--Clearcut, simple and mutually-exclusive category for those who identify as Native Hawaiians.

--Easy to ask in a telephone survey.

- Tabulation and analysis:

--American Indian groups, concerned with an accurate count of their population, preferred this option to including Native Hawaiians in the "American Indian or Alaskan Native" category.

--Provides specific information for policy development, trends analyses, needs assessments, program evaluation, health research, and civil rights enforcement.

Cons of Option (c)(4):

- Collection:

--Adds a category.

--High respondent error likely as some persons born in Hawaii but who do not have their origins in any of the original peoples of Hawaii likely to be confused by the term "Native." Because the Native Hawaiian population is relatively small (211,000 according to the 1990 census race question), respondent error could seriously affect the count.

--Unknown how data collected by visual observation would be affected.

--Addition of "Part-Hawaiian" will affect historical comparability of "White" category in Hawaii to some small extent.

- Tabulation and analysis: very small population group.

Past research results/literature review: The proportion of Asian and Pacific Islanders such as Cambodians and Laotians (groups not listed separately) reporting in the "other race" response circle to the 1990 census race item may be due to question design. Additionally, persons who were not Asians or Pacific Islanders marked the circle for "Other Asian or Pacific Islander." Of persons marking the "Other Asian or Pacific Islander" circle in the 1990 census, 54 percent of the write-ins were not consistent with the marked circle and nearly 40 percent were Hispanic group write-ins.

(d) American Indian or Alaska Native

The category of American Indian or Alaskan Native in Directive No. 15 includes persons having origins in any of the original peoples of North America and who maintain cultural identification through tribal affiliations or community recognition. This summary reports only on options proposed during public hearings and in the public comment period. It also highlights pros and cons for these options as raised in public comment or shown by research. Inclusion in the summary does not reflect OMB endorsement of the comments or suggestions. Requests included:

Options Suggested in Public Comments:

(1) Suggestions for change in category title include: "American Indian, Alaskan Indian, Eskimo, and Aleut"; "American Indian, Alaskan Indian, Aleut, or Eskimo"; "Federally Recognized American Indian and Alaskan Native"; and "Native American." Some prefer "Alaska Native" to "Alaskan Native." Suggestions also include collecting information on Tribal enrollment.

Pros of Option (d)(1):

- Collection: "Alaskan Indian," "Eskimo," and "Aleut" are more specific terms than "Alaskan Native" and reduce respondent error.

- Tabulation and analysis:

--"Federally recognized American Indian and Alaskan Native" meets Federal program needs as it indicates the legal (rather than racial) status of persons in this category. The federal trust responsibility to provide various educational, health, and housing services extends only to federally recognized American Indian and Alaskan Native tribes and their members and descendants of members. More people self-identify as being of American Indian or Alaskan Native race or descent than are enrolled in tribes or can prove descent, which tribal governments feel deprives their people of benefits rightfully belonging to them under Federal programs.

--Inclusion of the term, "Federally recognized" will affect historical continuity but for the future, it could clarify the intention of the category and reduce the changes over time in the numbers included in the category.

Cons of Option (d)(1):

- Collection:

--The term, "Alaskan Native" results in respondent error because some persons born in Alaska but who do not have Alaskan Indian, Eskimo, or Aleut origins are confused by the term.

--Some individuals of tribes not Federally recognized may not be aware of the status of their tribe (e.g., State recognized tribes or tribes awaiting recognition).

--When tribal enrollment/descent is not required information, possible overcoverage occurs because the category is marked by many persons with American Indian ancestry but no legal tribal affiliations or community recognition. This possible overcoverage could become more serious if there is an instruction

to "check all that apply" to allow multiracial persons to identify their specific racial groups and they respond in terms of ancestry further back than their immediate parents.

--The term, "Native American," is an unacceptable term to many American Indians. It is also confusing to some persons who are not American Indians but who use the term to indicate they were born in the United States. The term appears to include Native Hawaiians although this is not entirely clear.

"Aboriginal population," while technically correct, is considered by many to be a demeaning term.

"Indigenous populations" include persons having origins in any of the original peoples of North America, the Hawaiian Islands; American Samoa; Guam; and the Northern Marianas Islands. The terms, "aboriginal population," "indigenous populations," and "original peoples," are not generally understood and would likely result in misreporting.

--It is unclear where South American Indians, Russian and European aboriginal tribes, or Australian aborigines who have immigrated to the United States are classified. Some think the current Directive No. 15 categories exclude these populations. Others include in the definition of "American Indian," all the aboriginal peoples of North America (except Eskimos and Aleuts) and of Central and South America. Some suggest a separate category for "other indigenous tribes" to include tribes such as Mapuchi and Mayan.

(2) Change the category to include Native Hawaiians and other indigenous populations. Suggested category names include: "American Indian, Alaskan Native, or Native Hawaiian"; "American Indian, Alaskan Native, Native Hawaiian, and American Samoan"; "aboriginal population"; "indigenous populations"; and "Indigenous/Aboriginal People" (also see discussion under (c)(3) above).

Pros of Option (d)(2):

- Tabulation and analysis: Native Hawaiians are not Asians or immigrants to the United States.

Cons of Option (d)(2):

- Tabulation and analysis:

--There is a legal distinction between "American Indians and Alaskan Natives" and "Native Hawaiians." Native Hawaiians are not eligible for the majority of programs and services available to American Indians and Alaskan Natives. Indian tribes are self-governing political entities. For example, the legislative mandates for Indian Health Service and the Bureau of Indian Affairs pertain only to American Indians and Alaskan Natives. On the other hand, some programs for "Native Americans" includes Native Hawaiians

as well as American Indians and Alaska Natives. Native Hawaiians are of Polynesian/Pacific Islander descent and are not descendants of the original peoples of North America. They have a distinct culture and social environment. The category would be too heterogeneous for health research.

--Would affect historical continuity of the data: effect on the data for carrying out trust obligations toward American Indians and Alaskan Natives is unknown.

--Western Samoa is an independent nation and how to report could be a problem.

(3) Collect information on specific tribal affiliation and distinguish between Federally-recognized tribes and State-recognized tribes (Tribal affiliation is based on criteria established by the tribe, not self-identification.).

Pros of Option (d)(3):

- Tabulation and analysis: Meets Federal agency needs for policy development, trends analyses, needs assessments, and program evaluation and enforcement. A way to distinguish between legal and ancestral identification with the American Indian category.

Cons of Option (d)(3):

- Collection: Respondents may not know the difference between Federally-recognized and State-recognized tribes.
- Tabulation and analysis: --State-recognized tribal affiliation is not required for Federal purposes.

--Small numbers for most tribes would not provide meaningful statistics in surveys.

Past research results/literature review: Of persons reporting as "American Indian" in the 1990 census, 13 percent did not specify a tribe; this was an improvement from the 1980 census results. There was higher than expected growth rate of American Indians from 1980 to 1990 (as well as from 1970 to 1980) which raises questions about what the census race question is measuring for this population. Some of the change is attributed to growth and improvements in the census and outreach programs, some to misreporting (for example, some Asian Indian parents reported their children as American Indian), and some to shifts in self-identification from White to American Indian. The quality of the data for the American Indian population is of concern since it is a relatively small population (about 2 million in 1990) and the data are used to disburse Federal program funds to American Indian tribal and Alaska Native Village governments. About 2 million persons said they were American Indian in the race

question of the 1990 census; however, 8.7 million included American Indian in their response to the ancestry question.

(e) Multiracial

How to classify persons who identify with more than one race is perhaps the issue that has engendered the most controversy in the present review. For the most part, the public comment used the term, "multiracial" to refer to persons of two or more races. A variety of options were suggested in public comment for how to collect racial data from multiracial persons. They are shown below, followed by pros and cons cited for each option. Table 1 summarizes the options. This summary reports only on options proposed during public hearings and in the public comment period. It also highlights pros and cons for these options as raised in public comment or shown by research. Inclusion in the summary does not reflect OMB endorsement of the comments or suggestions.

In Latin America, a racially mixed society, there is an array of terms to describe gradations of skin color. This has not been the history of the United States in this century where the terminology implies "pure" races such as White or Black, rather than biracial or multiracial categories. In 1960, there were about 150,000 interracial marriages compared with 1.5 million in 1990. In the 1990 census, about 4 percent of couples reported they were of different races or one was of Hispanic origin. Such households had about 4 million children.

Directive No. 15 says that persons of mixed racial and ethnic origins should use the single category which most closely reflects the individual's recognition in his or her community. The public comments indicate that multiracial persons objected to this instruction. The commenters indicate that a single category does not reflect how they think of themselves. From their perspective, the instruction requires them to deny their full heritage and to choose between their parents. They feel they are being required to provide factually false information. They maintain that the current categories do not recognize their existence. They say they could mark "Other" where that category is provided but they feel it is demeaning. They want to identify their multiple races, but say that those who prefer to choose one of the existing broad categories could do so.

One concern of those who oppose a category for multiracial persons is that it will reduce the count for persons in the basic categories. Organizations representing multiracial persons disagree. They say minority groups could gain numbers as some persons are now classified as "White" under the "choose one" rule. As reflected in the options listed below, there was disagreement as to whether identification should include specific races. If specific races are identified, there might be some flexibility in how users could tabulate data. For some, this is seen as an advantage. For others, it is seen as a disadvantage because different tabulation rules would result in different counts of groups.

Some asked how far back in one's ancestry respondents should go in deciding to identify multiple races. Most who commented meant only the race or Hispanic origin of parents. This would require additional instructions and may not be acceptable to those who wish to identify their earlier ancestry. Presumably, persons would be instructed to list all races if the parent(s) were also of multiple races; this concerned those who oppose a multiracial category.

The discussion below refers to "race" but some respondents suggested multiple "ancestry" (listing both parents) should be the focus instead. Asking about ancestry focuses the questions back in time and conveys an historical and geographic context which some feel is clearer than the ambiguity of "race" or "ethnicity."

Table 1. Summary of Options for Identification of Multiracial Persons

(e)(1)	Multiracial identification not allowed (must pick one broad category):
(aa)	Individual chooses the one with which he or she most closely identifies
(bb)	Mother's category is designated
(cc)	Father's category is designated
(dd)	Race of minority-designated parent (if one is White)
(e)(2)	Multiracial identification allowed:
(aa)	"Multiracial" category -- self-identification (SI) or observer identification (OI)
(bb)	"Mark all that apply" from list of specific categories -- SI only
(cc)	Open-ended question -- SI or OI
(dd)	"Other" -- SI only
(ee)	Mother's and father's geographic ancestry -- SI only
(ff)	Skin-color gradient chart -- SI or OI

Options Suggested in Public Comments:

Option (e)(1): Mark one broad category with which the respondent most closely identifies (categories are same or similar to current list)

Pros to Option (e)(1) -- mark one broad category:

- Collection:

--Physical space on forms and questionnaires same as now.

--Identification: most people identify with only one of the current categories; facilitates collection by observers where that method is used; persons of multiracial heritage who identify with one broad category do not have difficulty responding.

--Telephone survey: easy to ask.

- Tabulations and analysis:

--Easier than options that allow the identification of multiple races.

--Meets needs of Federal agencies concerned with program evaluation and civil rights monitoring and enforcement.

--This is the only option that meets the needs of the Indian Health Service which is responsible for health care of anyone who is a Federally-recognized American Indian or Alaskan Native, regardless of the proportion of Indian blood or which parent has Indian blood.

--Maintains historical continuity of data.

--Categories are the same or similar to those used in State and local administrative records and historical Federal administrative records.

--Federal laws are written based on the assumption that people identify with one Directive No. 15 category. For civil rights monitoring and enforcement, respondents clearly fall in or out of a particular category. Would address concerns of those who believe a "multiracial" category would compromise effective implementation of civil rights laws.

Cons to Option (e)(1) -- mark one broad category:

- Collection -- identification and count issues:

--Having to choose one racial category upsets some respondents, especially those with immediate multiracial heritage, who identify with more than one race/ethnicity; telephone interviewers ask race in early part of interview and then must deal with an unhappy respondent for the remainder of the questions; and Federal agencies must respond to those upset by the policy.

--Nonresponse rates may increase for persons who wish to identify with more than one race but who are instructed to select the one category with which they most closely identify.

--As the size of the Hispanic population increases, a larger number and proportion of that population group may mark "Other" or not respond. The 1994 pretest of the Survey of Income and Program Participation showed some Hispanics would report in the multiracial category.

--Inconsistencies: The same person is likely to be identified differently across administrative records and surveys which reduces analytic usefulness of the data. If mother's race is used to assign the child's race as in birth records, the classifications may be different than the person's self-identification.

--There is a significant number of interracial marriages among Asians and Whites. For example, in the 1990 census, in California, nearly one-fourth of children with any Asian background were White and Asian. Asian groups contend they are undercounted when forced to identify with one category only. One study of the 1990 census indicates that the children of these marriages are more likely to identify themselves as "White" than as "Asian."

- Tabulations: Option (e)(1)(dd), in which the race of the minority-designated parent is designated as the person's category, requires additional rules if one parent is not White.

- Analysis:

--Does not sufficiently reflect Nation's diversity; no information for multiracial persons about differences in health, economic status, and likelihood of discrimination.

--Not as useful in health research as identification of the specific mixtures. All of the current racial categories are said to be too broad for analysis of health risks and economic trends; for example, a study found that 25 percent of those in the "Asian or Pacific Islander" category smoked, but this ranged from 20 percent of Filipinos to 72 percent of Laotians.

Option (e)(2)(aa): "Multiracial" category (SI or OI) (Note: May ask respondent to specify races but not necessarily)

Pros to Option (e)(2)(aa) -- "Multiracial" category:

- Collection if specific races are not identified:

--Physical space on forms: adds one racial category.

--Meets demand of some multiracial respondents, especially those whose parents are of different races.

--Telephone survey: easy to ask if it is the only category added; however, if additional categories are added may be problematic.

--Somewhat more amenable to identification by observers than any other option for multiracial persons (however, compared with observer identification in Option (e)(1), this option is likely to result in an undercount and a substantially different distribution of current broad categories).

- Tabulation and analysis:

--A few States have passed laws to include this category in their administrative records. Currently they proportion their multiracial counts among the OMB categories for Federal reporting purposes based on percentages of minorities in the general population, although it is not clear what geographic level they are using (National, State, local, school districts, etc.) when they refer to "general population." A change by OMB to a "multiracial" category would reduce costs for these few States because they would not have to maintain data in two different ways.

--Indication of population diversity.

--Potentially useful in analyzing trends such as education and employment, especially if specific categories are identified.

Cons to Option (e)(2)(aa) -- "Multiracial" category:

- Collection:

--Requires testing for effect on respondents, response rates, and data quality. Multiracial persons who previously identified principally with one broad category may become unsure of what is being asked. Multi-ethnic persons of the same race (e.g., a White person of English and German descent) may find the questions confusing. It is not clear how multiracial Hispanics would answer.

--The category is imprecise and specific instructions would be required on whether respondents should answer in terms of the races of their parents only or further back. Imprecision of the category leads to

possible confusion since, if one goes back far enough, many Americans are of mixed racial heritage (for example, many Whites have American Indian heritage and many Blacks have African, White, and American Indian heritage).

--Have to determine an acceptable category name. Suggestions included: Multiple races; Mixed race; Multiracial; Tan American; TIRAH (Tan InterRacial American Humankind); Mixed origins; Mestee; More than one race; and Mulatto. Some of these suggestions apply only to Black and White mixtures. Cognitive research shows that most people understand the terms, "multiracial" and "biracial."

--Requires establishment of a category (and the associated extra costs) but the category may be used by only a small proportion of the population. Some school systems allow use of a "multiracial" category and report it is used by less than 2 percent of students.

--Where identification is by an observer: unknown what criteria an observer would use to identify persons of more than one race; identification of specific races unlikely or too inaccurate to be useful; because of likely mismatch, unclear how it would affect a count of mixed race persons; broad category of blood relatives likely to be identified differently, especially in administrative records; and the same person is likely to be identified differently across administrative records and surveys which reduces the analytic usefulness of the category.

- Tabulations and analysis:

--The category is not an alternative in the administrative records of many State and local governments.

--If specific races are identified through an open-ended question, the development of a classification system for tabulating responses would be required; choices may be controversial and challenged.

--If specific races are not indicated, the general category is too heterogeneous for meaningful analysis or for use in civil rights monitoring and enforcement. A heterogeneous category does not provide sufficient information for health researchers (disease risk specific to racial and ethnic groups, monitoring of historical trends) and would complicate the design, conduct, and evaluation of health intervention programs. It is unclear how such a heterogeneous category could be used in civil rights monitoring and enforcement and such efforts could be more difficult and costly.

--There would be a major effect on historical continuity if specific races are not indicated because it may reduce the count of the current broad categories and in unknown ways. Some expressed concern that if

specific races are not known, the category has the potential for increasing racial segregation, discrimination, and the stigmatization of broad categories (other than White) which may result in less effective enforcement of civil rights laws.

--Persons with the same general cultural heritage and with similar physical characteristics may be classified differently.

Option (e)(2)(bb): "Mark all that apply" (SI only)

Pros of Option (e)(2)(bb) -- mark all that apply:

- Collection:

--If no new categories are added, physical space on forms and questionnaires same as now.

--Meets desire for self-identification of many multiracial respondents.

- Tabulation and analysis:

--Detail allows flexibility. Indicates extent and makeup of Nation's diversity.

--Can maintain some historical continuity by aggregating specific categories into current broad categories (for example, a person who has one Black parent and one White parent could be tabulated, depending on the purposes of the data, in three ways: White, Black, or Black/White). See related discussion below under "cons."

--Decision rules about aggregations of detailed categories could be discussed and documented.

Currently, we do not know what basis multiracial people use for marking their specific identity as a broad category.

--Provides potentially useful subgroup information for health researchers in terms of race-specific diseases, especially if the race of each parent is identified. For example, one study found a difference in the probability of low birth weight between Black mother-White father and White mother-Black father populations; small-for-gestational-age rates and preterm delivery rates also vary by race of the mother.

Cons of Option (e)(2)(bb) -- mark all that apply:

- Collection:

--Telephone survey: Difficult and may negatively affect data quality.

--In personal interviews, must use a flash card (can list responses on control card if survey will be done by telephone later); tedious in large households.

- Tabulations and analysis:

--Complex because of the many possible combinations. Historical continuity of counts and characteristics would be problematic. In the 1980 and 1990 censuses, for example, race was reported as "Black" for two-thirds of children in families with one Black parent and one White parent present. For families with a White parent and an Asian or Pacific Islander (API) parent, the proportion of children whose race was reported as "API" versus "White" was different in the last two censuses. Allocation rules would be controversial even if the objective is to achieve historical continuity to the extent possible (especially for characteristics).

--Aggregation decision rules would be required and may be subject to controversy.

--Current Federal laws are premised on persons identifying with one racial group. It is not clear what the impact would be for persons identifying as multiracial.

Option (e)(2)(cc): Open-ended question (SI or OI) (allows multiple responses)

Pros of Option (e)(2)(cc) -- open-ended question:

- Collection:

--Physical space on questionnaire/forms less than currently.

--Telephone survey: easy to ask if it is the only category added; however, if additional categories are added may be problematic.

--Meets desire for self-identification of many multiracial respondents, those who want to answer "American," and persons from small national-origin groups. Respondents likely to be satisfied since they are not restricted by pre-defined categories. One study of an open-ended question showed only 13

percent of Hispanic respondents used the conventional racial designations of "White" or "Black." For these Hispanics, self-identification was based more on cultural and ethnic identity. In the 1990 census, about 90 percent of the population reported an ancestry in the open-ended question; only 0.7 percent were uncodable responses; but about 10 percent did not respond to the question.

- Tabulations and analysis:

--Detail allows maximum flexibility and provides sociologically rich information for analyzing trends. Provides subgroup information useful to health researchers in terms of race-specific disease risk.

--Can maintain some historical continuity by aggregating specific categories into broad categories in 1977 Directive No. 15 (except see cons below and under (e)(2)(bb), "tabulations and analysis").

--Does not require respondent to indicate a preferred race; if desired, the tabulation rules can imply a priority by following the order of responses.

Cons of Option (e)(2)(cc) -- open-ended question:

- Collection:

--Unlikely States would collect data this way for their administrative records and thus, there would likely be a mismatch among data sets (also negative effect on analysis when trying to compare results among data sets).

--Same person likely to be identified differently across administrative records and surveys which reduces analytic usefulness.

--Does not allow for observer identification.

- Tabulations and analysis:

--Must develop a classification system to categorize hundreds of possible responses and the choices can be controversial. See discussion above in (e)(2)(bb) under cons, "tabulations and analysis."

--Negative effect on counts of broad categories and data quality, including considerations listed below:

(1) Religions given as responses cannot be tabulated into a broad category. This generates complaints (because of separation of church and state, religions cannot be tabulated by government agencies) and increases the effective nonresponse rate.

(2) National origins or nationalities are likely answers. Data collection agencies would have to code to broad categories based on probability (e.g., "English" likely to be White but could be Black or Asian also).

(3) "American" is a frequent response (the 6th ranked group in the 1990 census ancestry question with 12.4 million such responses or 5 percent of all responses) and cannot be coded to a broad category (effectively increases nonresponse rate). Foreign born and non-English speakers showed greater difficulty with open-ended write-in questions such as the 1990 census ancestry question.

--Negative effect on data quality: citing examples is interpreted as influencing respondents and giving no examples can also have significant effect on counts of broad categories. Research from the 1980 and 1990 censuses indicates high levels of inconsistent responses to open-ended questions and strong "example" effects. For instance, from 1980 to 1990, the number of Cajuns, which was an example in the ancestry question in 1990 but not in 1980, grew from 30,000 to 600,000. French, which was dropped as an example in 1990, declined from 13 million to 10 million.

--It is sometimes hard to interpret what respondents intend by their responses.

Option (e)(2)(dd): "Other -- specify" (SI) at end of list of broad categories

Pros of Option (e)(2)(dd) -- "other":

- Collection:

--Does not take up much physical space on the questionnaire.

--Telephone survey: easy to ask if it is the only category added; however, if additional categories are added may be problematic.

--Identification issues: respondents likely to find it easier to express their identity since they are not restricted to only the pre-defined categories (for example, those who want to answer "American" can do so); can ask multiracial respondents to choose one racial category but if they refuse, they can specify all the categories they choose; and allows identification of multiple and single categories not listed elsewhere

(e.g., Indians of South and Central American background).

- Tabulations and analysis:

--Detail allows maximum flexibility and provides sociologically rich information for analyzing trends. Potentially provides subgroup information useful to health researchers in terms of race-specific disease risk.

--Can maintain historical continuity by aggregating specific categories; however, see discussion in (e)(2)(bb) under cons, "tabulations and analysis."

Cons of Option (e)(2)(dd) -- "other":

- Collection:

--Some people are offended by identification as "Other"; when this proposal was made by OMB in 1988, it was not widely accepted and so was not adopted.

--Same person likely to be identified differently across administrative records and surveys which reduces analytic usefulness.

--If "Hispanic" is not listed as a separate category, research shows an extremely high percentage of "Other" responses are Hispanics who do not identify with one of the listed race categories.

- Tabulations and analysis:

--Must develop a classification system to categorize hundreds of possible responses and the choices can be controversial. See discussion in (e)(2)(bb) under cons, "tabulations and analysis."

--Detailed information not needed for program evaluation or civil rights monitoring and enforcement.

--If specific responses are not coded, the category is too heterogeneous to be useful.

--Negative effects on counts of broad categories and data quality, including considerations listed below:

(1) Religions given as responses cannot be tabulated into a broad category. This generates complaints

(because of separation of church and state, religions cannot be tabulated by government agencies) and increases the effective nonresponse rate.

(2) National origins or nationalities are likely answers. Data collection agencies would have to code to broad categories based on probability (e.g., "English" likely to be White but could be Black or Asian also).

(3) "American" is a frequent response (the 6th ranked group in the 1990 census ancestry question with 12.4 million such responses or 5 percent of all responses) and cannot be coded to a broad category (effectively increases nonresponse rate).

--Negative effect on data quality: citing examples is interpreted as influencing respondents and giving no examples can also have significant effect on counts of broad categories.

--It is sometimes hard to interpret what respondents intend by their responses.

Option (e)(2)(ee): Mother's and Father's Geographic Ancestry (SI only) (Respondent would be given a numbered geographic list and mark the appropriate numbers to indicate the region of origin of ancestors who migrated to the United States)

Pros of Option (e)(2)(ee) -- geographic ancestry:

- Collection:

--No overlapping categories.

--One clear concept, geographic origin of ancestors. In the 1990 census ancestry question, virtually all of the responses were national origin rather than ethnic origin (e.g., "Italian" more often than "Amalfi" or "Calabrian").

--Telephone survey: can do but not easily.

- Tabulations and analysis: Geographic origin may be a better indicator of health differences than race for many people. Tabulations would be lengthy but not difficult.

Cons of Option (e)(2)(ee) -- geographic ancestry:

- Collection

-- identification issues:

--People who can mark their race may not know the geographic region of origin of their ancestors or parents (e.g., adoptees).

--Likely to be a high rate of error in the "North America" category (only American Indians could correctly mark this category but it is likely those born in the United States would also mark it).

--Same person likely to be identified differently across administrative records and surveys which reduces analytic usefulness.

--Many people have ancestors from several different geographic regions and reports are often inconsistent among data sets.

- Collection

-- physical space on forms: considerably more than currently.

- Tabulation and analysis: Does not meet Federal program needs or provide historical continuity (for example, a person from Africa might be White or Black).

Option (e)(2)(ff): Skin-Color Gradient Chart (SI or OI) This is a suggestion for a numbered chart, a scale of skin-tone colors, reproduced on forms. Respondents would check the skin-tone number closest to the color of the individual respondent.

Pros of Option (e)(2)(ff) -- skin color chart:

- Collection: Less physical space on forms than now.

- Analysis:

--Can compare skin-tone responses with socioeconomic status and differences in effects of discrimination.

--A measure without racial labels less subject to changes in meaning over time as compared with labels based on race, ethnicity, and ancestry.

Cons of Option (e)(2)(ff) -- skin color chart:

- Collection:

--Identification: offensive to many; same person likely to be identified differently across administrative records and surveys which reduces analytic usefulness; individuals could change skin colors over a lifetime as a result of exposure to sunlight or disease.

--Telephone survey: impossible.

--Costly: requires precise, multicolor printing (color tones will vary among forms) when one-color (usually black) printing is now the case for most forms and questionnaires.

- Tabulation and analysis:

--No historical continuity; does not meet program needs.

--Skin color (melanin content) is not the sole way people identify their race and ethnicity. Culture, geography, and history, for example, are also considerations for many. For example, Black Africans and very dark Asian Indians may have similar skin tones but do not consider themselves in the same race category.

--Blood relatives may be coded differently.

--Not useful for health research or other types of socioeconomic research.

Past research results/literature review on a multiracial category: Some persons of mixed parentage or parents of interracial children who want to report more than one race are unsure how to respond. In the 1990 census, 98 percent of the population identified in one category; only 2 percent provided write-in multiple responses to the race question despite the instruction to mark one race only. Developing instructions for who should and who should not mark a "multiracial" category is difficult; in a 1994 pretest of the Census Bureau's redesigned Survey of Income and Program Participation, some persons thought they were being asked what race they would like to be if they could be multiracial even though their parents were from the same racial group.

(f) Hispanic origin

Directive No. 15 defines Hispanic as a person of Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish culture or origin, regardless of race. There is significant confusion in public comment as to whether Spaniards, Portuguese, Brazilians, and American Indians with a mixed heritage of Mexican or Central or South American tribes are included in the category, "Hispanic origin." Three major questions were raised. One is whether Hispanic origin should be a category in a single "race/ethnicity" question or whether there should be a question about Hispanic origin separate from race (discussed in Issue 3 above). The other two questions, on heterogeneity of the category and terminology, are discussed below. This summary reports only on options proposed during public hearings and in the public comment period. It also highlights pros and cons for these options as raised in public comment or shown by research. Inclusion in the summary does not reflect OMB endorsement of the comments or suggestions. Requests included:

Options Suggested in Public Comment:

(1) Collect data for population subgroups of the "Hispanic origin" category.

Pros of Option (f)(1):

- Tabulation and analysis: The category, "Hispanic origin," represents a heterogeneous population. Information on subgroups describes the significant social, economic, and health differences among the Puerto Rican, Mexican-American, Cuban, and other Hispanic populations.

Cons of Option (f)(1):

- Collection: Visual identification of nationality groups is difficult.

(2) Alternative or additional words suggested for "Hispanic" include "Latino/Hispanic Origin," "Latino," "Latin," "Latin American," and "Hispanics from the Americas" (to exclude persons from Spain and the Philippines). Persons of Mexican ancestry did not agree on terminology for their group. Some wanted "Pre-Columbian" because of their Mestizo (Indian) background. Others disagreed saying some Mexicans have European background. Some preferred the term, "Chicano" to identify Mexican-Americans while others found the term offensive.

Pros of Option (f)(2):

- Collection: Some respondents prefer an alternative.

Cons of Option (f)(2):

- Collection: The term, "Latino," includes a diverse group of people from many national origins, races, and backgrounds. Some understand the term, "Latin" or "Latino" to include Europeans such as Italians, French, Portuguese, Romanians, and Spaniards. Cognitive research by the Census Bureau indicates some understand "Latino" as meaning from Latin America, "Hispanic" as meaning someone who speaks Spanish, and "of Spanish origin" as someone from Spain or with a distant relative who was Hispanic.

Past research results/literature review: Results from the 1990 census showed that the Hispanic population of some 22.4 million grew by 53 percent from 1980 to 1990. Immigration accounted for about half the growth. Overall, the Census Bureau considers the quality of census and survey data for Hispanic origin to be good. Nevertheless, evaluations show high nonresponse (10 percent; research shows most are not Hispanics) and misreporting (for example, some non-Hispanics report in the "Mexican-Amer." category to indicate they are American). In the 1990 census race question, two in three persons who did not mark a race circle, wrote in a response reflecting Hispanic ethnicity. Among persons who indicated in the 1990 census that they were of Hispanic origin, 52 percent marked the "White" circle and 43 percent marked the "Other race" circle. Based on evaluations of the 1980 Census and 1990 Census pretests, it appears that persons reporting "Other Spanish/Hispanic," included Brazilians and other persons of Portuguese descent who feel the term, "Hispanic," also applies to them.

C. Future Research Agenda

Agency staff and funding for research and testing associated with possible changes are very limited. As a result, plans necessarily have to be developed within those resource constraints and may change. Within available resources, Federal agencies are conducting research through 1996 to inform decisions on selected options. A brief summary of the future research agenda, as of April 1995, is presented in this section. The number of issues that can be tested in 1995 and 1996 is limited. This *Federal Register* notice provides the last opportunity for public comment on priorities for research in 1996.

Research Agenda

The Interagency Committee's Research Working Group, which is co-chaired by the Bureau of the Census and the Bureau of Labor Statistics, reviewed all the criticisms and suggestions for changing the current categories that appeared in OMB's June 9, 1994, *Federal Register* notice, including requests received during the public comment period to expand the standards by establishing additional categories for specific population groups. Some of the more significant issues that have been identified for research and testing are: classification of multiracial persons; combining race and Hispanic origin; combining concepts of race/ethnicity/ancestry; changing the names of current categories; and adding new classifications. The Race and Ethnic Targeted Test, to be conducted by the Bureau of the Census in 1996, will be the major opportunity to test three to four options on race and ethnicity.

The Bureau of Labor Statistics designed a Supplement to the May 1995 Current Population Survey (CPS) to provide information about three issues with respect to Directive No. 15. They are (1) what proportion of respondents will choose a "multiracial" category and how that may impact on the data for the other racial categories; (2) inclusion of an Hispanic category in the list of races; and (3) preferences concerning specific terms such as "African American" and "Latino." To gather this information, the Supplement is divided into four panels, and a random sample of approximately 15,000 of the 60,000 CPS households will receive one of the following four survey instruments.

Panel 1: Separate race and Hispanic origin questions; no multiracial category

Panel 2: Separate race and Hispanic origin questions; with a multiracial category and races specified

Panel 3: A combined race and Hispanic origin question; no multiracial category

Panel 4: A combined race and Hispanic origin question; with a multiracial category and races specified

In addition, all households in the May Supplement will be asked questions about their ancestry, preferences concerning specific terms, and use of languages other than English in the home. The ancestry and language questions are included to help explain differences in reporting by households with similar racial characteristics. Results of this test are expected to be available in late Fall 1995.

Multiracial Category.-- Research and testing of a multiracial category is especially important since it could have a significant impact on the usefulness of data resulting from the current racial and ethnic categories. An important aspect of this issue on which research needs to be conducted is the extent to which persons of mixed racial heritage will identify in a separate multiracial category on surveys and censuses.

To begin research on this issue, a multiracial response option was included in operational pretests for the revised Survey of Income and Program Participation involving 292 households in the Atlanta, Boston, and Chicago metropolitan areas during April and May 1994. Despite the small sample size, the results were somewhat informative for two reasons: (1) a higher percentage (7.3 percent) of persons reported in the multiracial category than have done so in some of the records from school and military systems cited in various public hearings and conferences, and (2) in nearly two-thirds (65 percent) of the 55 write-ins to the multiracial item, the respondent reported as Hispanic (23 cases or 42 percent) or as Hispanic and some other race group. The higher percentage reporting as multiracial might reflect the sites of the pretest and the oversampling of low and high income areas. The high proportion of multiracial responses involving Hispanics does indicate that a multiracial category might draw disproportionately more responses from Hispanics than from the other racially mixed persons for whom many were seeking this option. These results underscored the importance of testing the multiracial category in larger samples (as in the May 1995 CPS Supplement), as well as perhaps the need for

additional definitions or instructions for the category if the intention is to draw responses primarily from persons whose parents are of different races. These early findings also served to indicate that cognitive research would aid in developing that Supplement.

In preparation for the May 1995 CPS Supplement, cognitive research interviews were conducted in 1994 and early 1995 with individuals who have parents of different races, as well as individuals who may identify with only one race, even though they may have a mixed heritage. The main objective of this cognitive research was to examine how individuals view race and ethnicity and how they might interpret and respond to a race question that provides a "multiracial, specify" option.

Combining Race and Hispanic Origin. -- The May 1995 CPS Supplement will provide needed research on whether a combined race/Hispanic ethnicity question should be used instead of separate questions on race and Hispanic ethnicity. Important reasons to research this issue are that some Federal agencies have been collecting and reporting data in a combined format for a number of years, and a high percentage of Hispanics selected "other race" in the 1990 decennial census race question when race and ethnicity were collected in two separate questions. Research questions include examining the effects of having a single race and Hispanic ethnicity question on the counts for other races and for Hispanics; examining which subgroups to include as "Hispanic"; determining what percentage of administrative record data bases already use "Hispanic" as a racial category and what percentage of respondents in these data bases are missing information on Hispanic ethnicity; and deciding if Hispanic ethnicity should be assumed to take priority over other racial categories (e.g., Black Hispanics).

In considering this issue, one should bear in mind that the concepts of race, ethnicity, and ancestry are not clearly or consistently distinguished in the U.S. population. For example, some Hispanics regard the "Hispanic" designation as a "racial" category, defining "race" in terms of national origin and cultural characteristics. As discussed below, it has been suggested, therefore, that census and survey respondents be asked about only a single concept -- perhaps ethnicity or race/ethnicity -- corresponding to self-perceived membership in population groups that might define themselves by cultural heritage, language, physical appearance, behavior, or other characteristics.

Combining Concepts of Race/Ethnicity/Ancestry.--Directive No. 15 has been criticized for not clearly distinguishing among race, ethnicity, and ancestry. Directive No. 15 specifically notes the absence of anthropological or other scientific bases for their separate designation. Varied and possibly inconsistent definitional criteria, such as geographic origin, cultural origin, cultural identification and affiliation, community recognition, and race itself, are used to describe the terms. The current Federal categories have created five single aggregations from heterogeneous and highly diverse populations. Since ethnic groups evolve and may change their group name over time, research is needed on the basic concepts to be measured as well as on the popular terminology respondents use to refer to their ethnic group. This research will be helpful in determining those response categories which would provide useful information about our Nation's population.

The research on this issue needs to consider a number of implications of combining the concepts. The consolidation of questions of "race," "ethnicity," and "ancestry" into a single question of "ethnicity" (or "race/ethnicity") or of "identified population groups" would eliminate the distinction between race and ethnicity indicated in Directive No. 15. Consolidation of the categories would also address the issue of including Hispanics as a racial designation rather than as a separate ethnic category. Under consolidation, Hispanic would be included as an ethnic or racial/ethnic category along with other categories previously classified as races. If, in addition to consolidating categories, respondents are allowed to select more than one ethnic or racial/ethnic identity, the issue of "multiracial" identification might also be addressed. The combined question would most likely solicit multi-ethnic as well as multiracial responses. In the 1990 census ancestry question, which allows multiple reporting of ethnicities, about 30 percent of the population reported multiple ancestries. Such a large proportion of multiple responses would present processing problems for Federal agencies. The consolidation of race and ethnicity would interrupt the continuity of categorization in the race and ethnicity questions in recent decades; however, continuity is already imperfect due to changes in questions and response options.

Terminology for Categories.--This issue is concerned with whether to replace or revise current terminology for Black, Hispanic, or American Indian racial/ethnic categories for data collection and data reporting with terms that have been suggested such as African American, Latino/Latina, and Native American. Research is needed to determine whether, and in what ways, any proposed changes in terminology may affect reporting or data collection. If a change in terms produces a change in coverage, it is useful to know what that change signifies. Any replacement of terminology should consider: (1) that the new terms might have meanings different from the old terms for respondents while, for the users, the old and new categories might appear synonymous; (2) that as current usage changes, terms are likely to have different meanings to people, and the new terms may exclude persons who were comfortable with the old terms but who may not perceive themselves as "fitting" under the new designation; and (3) the extent to which definitions need to accompany new categories. Questions about preferences for various terms are included on the May 1995 CPS supplement.

Additional research plans:

- The Census Bureau is conducting cognitive research from February through July 1995 on issues such as a multiracial category, marking all categories that apply, terminology, and a combined race/Hispanic origin/ancestry question. Research on the classification of "Native Hawaiian" is also planned. The extent of research is dependent upon available resources. The Census Bureau also plans to conduct two tests in 1996: the National Content Test (NCT) and the Race and Ethnic Targeted Test (RAETT). The NCT is designed to test selected population and housing questions for the 2000 census. It will be a national sample of 35,000 to 50,000 households. To determine what information respondents will provide in a self-reporting context, the Census Bureau has identified a multiracial category or response option (for

example, multiple responses) as a high priority for panels on the 1996 National Content Test. Other issues to be tested in the NCT include terminology and the placement of the Hispanic origin question first, followed immediately by the race question. The RAETT, which will include a reinterview, will provide the most extensive opportunity to test several options for collecting racial and ethnic data. The proposed sample of about 90,000 households will be targeted to a diverse sample of racial and ethnic populations. The Census Bureau expects that the RAETT will allow further testing of a multiracial classification, terminology, and other selected options.

- The National Center for Health Statistics and the Office of the Assistant Secretary for Health will conduct research on the effects of changes in racial classification on birth certificate records.
- The Centers for Disease Control and Prevention is undertaking a project to evaluate the recording of racial classifications on death certificates. This study will involve a survey of a sample of funeral directors with the aim of improving the quality of racial data reported on death certificates.
- A literature search on work related to racial classification in the health field (using MEDLINE) is being conducted by the Department of Health and Human Services (DHHS).
- An inventory of DHHS minority health data bases is being developed by the DHHS. It will provide information on what data are available and data collection problems that have been encountered.
- The National Center for Education Statistics (NCES) is conducting a Spring 1995 survey to obtain information: (a) how schools currently students' collect racial and ethnic data; (b) how administrative records containing racial and ethnic data are maintained and reported; (c) what State laws mandate or require of school systems with respect to collecting data on race and ethnicity; and (d) current issues in schools regarding race and ethnicity categories.

D. General Principles for the Review of the Racial and Ethnic Categories

The criticisms and suggestions for changing Directive No. 15 have underscored the importance of having a set of general principles to govern the current review process. The following principles were drafted in cooperation with Federal agencies serving on the Interagency Committee. The principles listed below are those OMB may use to guide final decisions on standards for the classification of racial and ethnic data. The principles are, for the most part, the same as those published in the June 9, 1994, *Federal Register* notice. There are changes to Principles 2, 5, 6, and 8. Principles 12 and 13 are new. The public is invited to comment on these or suggest additional principles.

1. The racial and ethnic categories set forth in the standard should not be interpreted as being primarily biological or genetic in reference. Race and ethnicity may be thought of in terms of social and cultural characteristics as well as ancestry.
2. Respect for individual dignity should guide the processes and methods for collecting data on race and ethnicity; ideally, respondent self-identification should be facilitated to the greatest extent possible, recognizing that in some data collection systems observer identification is more practical.
3. To the extent practicable, the concepts and terminology should reflect clear and generally understood definitions that can achieve broad public acceptance. To assure they are reliable, meaningful, and understood by respondents and observers, the racial and ethnic categories set forth in the standard should be developed using appropriate scientific methodologies, including the social sciences.
4. The racial and ethnic categories should be comprehensive in coverage and produce compatible, nonduplicated, exchangeable data across Federal agencies.
5. Foremost consideration should be given to data aggregations by race and ethnicity that are useful for statistical analysis and program administration and assessment, bearing in mind that the standards are not intended to be used to establish eligibility for participation in any Federal program.
6. The standards should be developed to meet, at a minimum, Federal legislative and programmatic requirements. Consideration should also be given to needs at the State and local government levels, including American Indian tribal and Alaska Native village governments, as well as to general societal needs for these data.
7. The categories should set forth a minimum standard; additional categories should be permitted provided they can be aggregated to the standard categories. The number of standard categories should be kept to a manageable size, as determined by statistical concerns and data needs.
8. A revised set of categories should be operationally feasible in terms of burden placed upon respondents; public and private costs to implement the revisions should be a factor in the decision.
9. Any changes in the categories should be based on sound methodological research and should include evaluations of the impact of any changes not only on the usefulness of the resulting data but also on the comparability of any new categories with the existing ones.

10. Any revision to the categories should provide for a crosswalk at the time of adoption between the old and the new categories so that historical data series can be statistically adjusted and comparisons can be made.
11. Because of the many and varied needs and strong interdependence of Federal agencies for racial and ethnic data, any changes to the existing categories should be the product of an interagency collaborative effort.
12. Time will be allowed to phase in any new categories. Agencies will not be required to update historical records.
13. The new directive should be applicable throughout the U.S. Federal statistical system. The standard or standards must be usable for the decennial census, current surveys, and administrative records, including those using observer identification.

The agencies recognize that these principles may in some cases represent competing goals for the standard. Through the review process, it will be necessary to balance statistical issues, needs for data, and social concerns. The application of these principles to guide the review and possible revision of the standard ultimately should result in consistent, publicly accepted data on race and ethnicity that will meet the needs of the government and the public while recognizing the diversity of the population and respecting the individual's dignity.

Sally Katzen

Administrator,

Office of Information and Regulatory Affairs

APPENDIX

DIRECTIVE NO. 15

RACE AND ETHNIC STANDARDS FOR FEDERAL STATISTICS AND ADMINISTRATIVE REPORTING

(as adopted on May 12, 1977)

This Directive provides standard classifications for recordkeeping, collection, and presentation of data on race and ethnicity in Federal program administrative reporting and statistical activities. These classifications should not be interpreted as being scientific or anthropological in nature, nor should they be viewed as determinants of eligibility for participation in any Federal program. They have been developed in response to needs expressed by both the executive branch and the Congress to provide for the collection and use of compatible, nonduplicated, exchangeable racial and ethnic data by Federal agencies.

1. Definitions

The basic racial and ethnic categories for Federal statistics and program administrative reporting are defined as follows:

1. **American Indian or Alaskan Native.** A person having origins in any of the original peoples of North America, and who maintains cultural identification through tribal affiliations or community recognition.
2. **Asian or Pacific Islander.** A person having origins in any of the original peoples of the Far East, Southeast Asia, the Indian subcontinent, or the Pacific Islands. This area includes, for example, China, India, Japan, Korea, the Philippine Islands, and Samoa.
3. **Black.** A person having origins in any of the black racial groups of Africa.
4. **Hispanic.** A person of Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or origin, regardless of race.
5. **White.** A person having origins in any of the original peoples of Europe, North Africa, or the Middle East.

2. Utilization for Recordkeeping and Reporting

To provide flexibility, it is preferable to collect data on race and ethnicity separately. If separate race and ethnic categories are used, the minimum designations are:

a. Race:

- American Indian or Alaskan Native
- Asian or Pacific Islander
- Black
- White

b. Ethnicity:

- Hispanic origin
- Not of Hispanic origin

When race and ethnicity are collected separately, the number of White and Black persons who are Hispanic must be identifiable, and capable of being reported in that category.

If a combined format is used to collect racial and ethnic data, the minimum acceptable categories are:

- American Indian or Alaskan Native
Asian or Pacific Islander
Black, not of Hispanic origin
Hispanic
White, not of Hispanic origin.

The category which most closely reflects the individual's recognition in his community should be used for purposes of reporting on persons who are of mixed racial and/or ethnic origins.

In no case should the provisions of this Directive be construed to limit the collection of data to the categories described above. However, any reporting required which uses more detail shall be organized in such a way that the additional categories can be aggregated into these basic racial/ethnic categories.

The minimum standard collection categories shall be utilized for reporting as follows:

1. **Civil rights compliance reporting.** The categories specified above will be used by all agencies in either the separate or combined format for civil rights compliance reporting and equal employment reporting for both the public and private sectors and for all levels of government. Any variation requiring less detailed data or data which cannot be aggregated into the basic categories will have to be specifically approved by the Office of Management and Budget (OMB) for executive agencies. More detailed reporting which can be aggregated to the basic categories may be used at the agencies' discretion.
2. **General program administrative and grant reporting.** Whenever an agency subject to this Directive issues new or revised administrative reporting or recordkeeping requirements which include racial or ethnic data, the agency will use the race/ethnic categories described above. A variance can be specifically requested from OMB, but such a variance will be granted only if the agency can demonstrate that it is not reasonable for the primary reporter to determine the racial or ethnic background in terms of the specified categories, and that such determination is not critical to the administration of the program in question, or if the specific program is directed to only one or a limited number of race/ethnic groups, e.g., Indian tribal activities.
3. **Statistical reporting.** The categories described in this Directive will be used at a minimum for federally sponsored statistical data collection where race and/or ethnicity is required, 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. A repetitive survey shall be deemed to have an adequate sample size if the racial and ethnic data can be reliably aggregated on a biennial basis. Any other variation will have to be specifically authorized by OMB through the reports clearance process. In those cases where the data collection is not subject to the reports clearance process, a direct request for a variance should be made to OMB.

3. Effective Date

The provisions of this Directive are effective immediately for all new and revised recordkeeping or reporting requirements containing racial and/or ethnic information. All existing recordkeeping or reporting requirements shall be made consistent with this Directive at the time they are submitted for extension, or not later than January 1, 1980.

4. Presentation of Race/Ethnic Data

Displays of racial and ethnic compliance and statistical data will use the category designations listed above. The designation "nonwhite" is not acceptable for use in the presentation of Federal Government data. It is not to be used in any publication of compliance or statistical data or in the text of any compliance or statistical report.

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

(1) The designations "Black 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;

(2) The designations "White," "Black," and "All Other Races" when the distinction among the majority race, the principal minority race and other races is appropriate; or

(3) The designation of a particular minority race or races, and the inclusion of "Whites" with "All Other Races," if such a collective description is appropriate.

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

When the primary focus of a statistical 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.