Spotlight on Heterogeneity

The Federal Standards for Racial and Ethnic Classification

Summary of a Workshop

Barry Edmonston, Joshua Goldstein, Juanita Tamayo Lott,
Editors

Committee on National Statistics
Commission on Behavioral and Social Sciences and Education
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1996

CLASSIFICATION PROBLEMS UNDER DIRECTIVE 15

Several drawbacks of the current directive were pointed out by federal agencies and other workshop participants. Many of these drawbacks can be grouped into two major areas. First, Directive 15 has been in some respects too successful in universalizing the five race and ethnicity categories. There has been little research about the substantive meaning or relevance of the categories, and most agencies do not evaluate the racial and ethnic data they collect. The specification of reporting categories has driven the construction of response categories, with agencies, except for the Census Bureau, deferring to the default categories, instead of constructing finer classifications that might be tailored to more specific needs.

Second, the current categories and the reliance on self-identification pose problems for many individuals and groups. Some respondents do not identify with or find applicable any of the available categories, while others encounter technical difficulties resulting from the current wording of the directive. For example, persons of Hispanic ethnicity are explicitly assumed to be either white or black and not Asian and Pacific Islander or American Indian and Alaskan Native.[5] There is no race category that includes persons native to Central and South America. There is no race category for blacks who come from areas in the world other than Africa, and there is uncertainty about many persons from northern parts of Africa. Other examples abound. The classification of Brazilians in the current directive is problematic, since they are South American but are not of "Spanish" culture or origin. Similarly, persons whose nation of origin is Spain are separated from others of European nationality.

Inconsistent Categories and Data Sources

Standardization of categories is of great importance to agencies and researchers calculating rates of any kind because rates require compatible data for numerator and denominator. Often, however, numerator and denominator data come from different sources. Workshop participants stated that such data are drawn from both federal and nonfederal agencies, as well as other sources. For example, data for demographic estimates at the national level come from vital statistics collected by individual states, which may use different race and ethnicity categories from Directive 15 and from each other. Often, the census is used as the principal source of numbers for the denominators, which can lead to systematically poor estimation if the collection procedures used for the numerator data

[5]Data from recent censuses shows that the latter category of Hispanics with American Indian ancestry is quite large: in 1990, 8.4 percent of people who self-identified as American Indian by race also checked Hispanic origin.

(e.g., birth or death certificate) do not gather data in the same manner as the census.

Infant mortality rates by race and ethnicity are the most frequently cited example of inaccuracies that result from inconsistent racial and ethnic reporting. Beginning in 1989, births have been categorized by the race of the mother, as recorded on the birth certificate.[6] If the mother does not state her race on the certificate, then the baby's race is imputed to be the race of the father. If neither parent's race is reported on the birth certificate, the baby's race is imputed to be the race of the mother on the preceding record with known race (National Center for Health Statistics, 1995a:54).

There were several reasons for this change in policy. First, U.S. Standard Certificates of Live Birth were modified in 1989 to include more questions that were directly associated with the mother's health and health behaviors. Moreover, because of increased incidence of interracial births and nonmarital births, it was determined that "tabulating all Births by the race of the mother provides a more uniform approach rather than a necessarily arbitrary combination of parental races" (National Center for Health Statistics, 1995a:54).

Racial classification of deaths, unlike births, is designated by an outside party, usually attending physicians or funeral home directors, on the basis of information provided by relatives or their own observations. Their reliance on visual identification or possibly inadequate information from relatives has resulted in an overassignment of deaths to black and white categories and an underassignment of deaths to American Indian and Asian and Pacific Islander categories (Hahn, 1993). As shown in Table 4, with the pre-1989 algorithm for parents of the same race, almost 10 percent of American Indian infant deaths were incorrectly classified even when both parents were American Indians, and the inconsistent assignment of Asian infant deaths varied from almost 20 percent for Japanese to almost 40 percent for Chinese. Changes introduced with the post-1989 algorithm, whereby infants are assigned the same race as their mothers, have failed to reduce the level of incorrect assignment. As a result, the reported infant mortality rates—even when both parents are Asian or both are American Indian—are substantially lower than the actual levels. Table 4 shows that the level of inconsistent classification of infant deaths by race is even greater for children of mixed heritage.

The result of this inconsistent reporting is only a slight overreporting of black and white infant mortality rates, since the numbers of births are so large.

[6]Prior to 1989, there were a complicated set of rules for classifying the race of a newborn. Newborns with two parents of the same race were classified as that race; newborns with one nonwhite parent were classified as the race of the nonwhite parent. When both parents were nonwhite but of different races, the newborn was assigned the father's race; except that if either parent was Hawaiian, the newborn was classified as Hawaiian (National Center for Health Statistics, 1995a).

TABLE 4 Proportion of Infant Deaths with Race Reported at Death Different from that Reported at Birth, 1983-1985

Infant Race at Birth	Parents' Races as Indicated on Infant's Birth Certificate: Same	Different
Pre-1989 Algorithm[a]		
All Races Combined	1.8	59.5
White	0.8	—
Black	3.0	39.7
American Indian	9.8	76.4
Chinese	38.6	68.2
Japanese	19.4	75.0
Hawaiian	3.5	33.8
Filipino	27.8	80.4
Other Asian	29.5	81.7
Post-1989 Algorithm: Newborns Classified as Race of Mother		
All Races Combined	1.8	56.7
White	0.8	45.5
Black	3.0	37.9
American Indian	9.7	71.2
Chinese	38.6	78.3
Japanese	19.4	86.7
Hawaiian	3.5	32.4
Filipino	27.8	85.1
Other Asian	29.5	84.3

NOTE: For example, of every 100 deaths of infants who were classified as "Other Asian" at birth and whose parents had different races (from each other), 81.7 percent were classified differently at death.

[a]See text for discussion.

Source: Adapted from Hahn (1993)

For the smaller racial and ethnic groups, however, the underreporting of infant mortality is quite large. Race- and ethnicity-specific life-expectancy calculations, relying on these infant mortality rates, will produce inaccurate results and have major ramifications for calculations made by Social Security and private insurance industry actuaries, among others.

It is important to understand that the inaccuracies in race- and ethnicity-specific infant death rates stem from a combination of three factors. The first is the difference in the reporting source for births and deaths. This is not a product of the formulation of Directive 15, but rather of the reporting processes used by the state and local authorities who produce the nation's vital statistics. It might be possible to improve consistency in racial and ethnic reporting, but it would

probably involve very detailed rules, and it is not clear that extensive guidance should be part of a standard. The only way to ensure that the reported race on a person's death or birth certificate is the same as in the census would be to have some kind of requirement for consistent individual identification. Such identification, tantamount to an individual race identification system, would be antithetical to the nation's history of individual privacy and would also contradict current understanding about the fluidity of racial and ethnic identity.

The second factor in the inaccuracies is the great inequality in size between groups. Since infant deaths are relatively rare events—about 1 percent of all births—even minor inconsistencies in the classification of deaths may have a major impact on the rates.

The third factor is that intermarriage is more common for American Indians, Asians, and Hispanics than for blacks and whites; hence, a higher proportion of children with an American Indian, Asian, or Hispanic mother have a father of a different ancestry (usually white). For these newborns, given the simple categories of Directive 15, there is necessarily inaccuracy in reported race and ethnicity. One can consider the operation of all three factors, for example, for mothers who are American Indians: small numbers of births, often born to mixed ancestry parents, and shifting racial identity over the lifetime. Calculation of mortality becomes extremely problematic in this situation.

Overlap of Race and Ethnicity Questions

A major problem in the racial and ethnic classification system is the inconsistency regarding Hispanic origin in the two formats permitted by Directive 15. The line between ethnicity and race is conceptually difficult. The directive attempted to provide a compromise, enabling data users and data collectors to choose the best-suited question and presentation formats for their purposes. However, the two formats present serious problems of incomparability and inconsistency.

According to workshop presentations, the incomparability of the two question formats appears in two ways. First, although the one-question format addresses the issue of black and white Hispanics, it does not allow the existence of Hispanics who also have American Indian or Asian and Pacific Islander heritage. Possible responses in the two-question format are not even included in the one-question format. Second, by making Hispanic origin mutually exclusive of all four race categories, the single-question format defines white, black, American Indian, and Asian and Pacific Islanders differently than does the two-question format.

The ramifications of such an inconsistency are not only theoretical, but can also have important quantitative impacts. For example, in his review of various combinations of race and Hispanic-origin categories based on the March 1991 Current Population Survey (CPS), del Pinal (1992:3) found that Hispanics ac-

Revising Federal Standards: Issues for Consideration

The U.S. Office of Management and Budget is considering revision of the federal standards of racial and ethnic classification in response to two interrelated factors: the demographic and social changes that are occurring in the United States and increasing expressions of dissatisfaction with the current standards among data users, data providers, and the public. Moreover, these two factors affect and are affected by changes in people's subjective views of their own racial and ethnic identity.

From a demographic perspective, the U.S. population has reached a stage at which it is ethnically diverse and ethnic intermarriage is increasingly common; consequently, there will be increasing numbers of people with multiple ancestries, for whom future preferences for self-identification are unknown. Workshop participants observed that these factors raise questions about the usefulness of demographic analysis and population projections based on conventional assumptions of "closed" ethnic groups with no exogamy. Perhaps the largest degree of consensus at the workshop was that any revision in the standard will itself need to be able to adapt to change.

Yet this recognition of the need for flexibility and adaptation conflicts with other considerations. For example, the responses by federal agencies to questions of criteria for standards showed fairly strong agreement on three issues (see below): exhaustive and mutually exclusive categories; simplicity in categorization; and as much continuity with historical data as possible.[1] Agencies also

[1]The 1970 census definition of Hispanic, for example, based on surname allocation, makes it virtually impossible to compare data from it with data from the 1980 and 1990 censuses, based on self-identification.

agreed that considerable field testing should be undertaken before changes are made. It is not possible to both meet these criteria and have a flexible and adaptable system that recognizes a variety of subjective self-perceptions of racial and ethnic identity.

This conflict about the criteria for a classification standard relates in part to the many purposes for the standard, which are far more varied than originally intended. At the federal level, the standards are used for statistical records and analysis, program administration, and civil rights compliance. Although Directive 15 was originally promulgated solely for the use of federal agencies, it has become the de facto standard for state and local agencies, the private sector, the nonprofit sector, and the research community. Users of the data are limited by available reporting and tabulations unless they go back to original survey tapes and create their own cross-classification tables by other groupings. In many cases even this is impossible, since the responses were coded in accordance with the Directive 15 categories.

Directive 15 has also had a trickle-down effect. Since state and local agencies and the private and nonprofit sectors often use Directive 15 for information they report to the federal government, they may use the same categories to achieve internal consistency, even for data that are not destined for federal reporting. Similarly, researchers often use the Directive 15 race and ethnicity categories because they are the only classifications available in the data.[2] Although Directive 15 was never intended to establish a national standard for race categories, it has come to function partly in that way. Thus, in considering revising federal standards, it is important to keep in mind that they are used to meet a wide range of data needs, collection methods, and presentation formats, from the detailed tables in government publications to the question format given to school children for enrollment.

BASIS FOR CLASSIFICATION

The Directive 15 standards for reporting of race and ethnicity are explicitly nonscientific, stating: "These classifications should not be interpreted as being scientific or anthropological in nature . . ." Although the directive is clear about what it is not, it is silent on the basis of its chosen categorization. National origin, race, culture, and community recognition are all combined to create a complicated and inconsistent classification scheme. At the workshop, one frequently heard suggestion for improving the categorization would be to state more clearly the principles on which it is based.

---

[2]The major exception is research analysis of decennial census data, for which specific ethnic or ancestry categories are often created for analysis. And as noted above, some researchers also conduct their own surveys and develop their own race and ethnicity coding schemes.

One of the prime criticisms of the current classification is that it lacks a consistent logic. Some of the categories are racial, some are geographic, some are cultural. Several participants commented that a primary concern of respondents, if not of federal agencies, is the perception of a fair treatment in the definition of categories. The categories now represent a combination of historical, legal, and sociological factors, but that is not explicitly acknowledged.

Much of the difficulty in the creation of a consistent, rational classification system lies in the fluid nature of what race and ethnicity are. Race and ethnicity are inherently complex concepts, with multiple sources of definition. There is no scientific basis for the legitimacy of race or ethnicity as taxonomic categories. That is, although there clearly are many and varied racial and ethnic distinctions, their multiplicity of sources defies a single-variable classification scheme based on a single individual characteristic. The challenge of creating logically consistent standards is magnified even more by self-definitions of race and ethnicity, in which a devised set of categories may well not coincide with people's views of themselves.

Schematization of the current classification highlights areas of inconsistency in the current definitions (see Table 2, above, and Appendix B). Geographic origin is the only criterion applied to all five categories, although the term used is "a person having origins in," and "origin" is not defined. Since the initial origins of all humans are unknown, the term does not refer to the far distant past; nor does it refer to the immediate past, when the origins of all U.S. residents except immigrants would be North America. Rather, the concept is indeterminate as to timing—one's parents? grandparents? any ancestor?—and so very open to differences in interpretation, especially if the basis of classification is self-identification.

The black category is the only explicitly racial one: "A person having origins in any of the black racial groups of Africa." It is also inconsistent with the definition for the white category: "A person having origins in any of the original peoples of Europe, North Africa, or the Middle East," which does not include anything about racial membership.[3] The racial definition for black is apparently meant to exclude white immigrants from Africa. Cultural affinity is used to define Hispanics and American Indians, although in slightly different ways. Hispanics are "of" Spanish culture, while American Indians or Alaskan Natives "maintain cultural identification through tribal affiliation or community recognition."

From a purely taxonomic standpoint, the current classification suffers from several faults. It does not cover persons of Australian or New Zealand origin. It places persons from Spain in both the white and Hispanic categories, since Spain

[3]The lack of a definition of "origin" is particularly noticeable for this category, which uses both "having origins in" and "original people of."

is a European country. It allows no place for American Indians whose cultural identity is not recognized by a tribe, but who are nonetheless descendants of the original peoples of North America. It leaves ambiguous or confused the status of many groups, including people from Brazil, Madagascar, and Cape Verde. From a social standpoint, the current definitions seem to reinforce the link between dark skin color and race, while ignoring skin color in the definitions of nonblack groups.

Although some observers protest the inconsistency and conceptual vagueness of the current categories, defenders of them point out that the very complexity of the definitions reflect the real-world complexity of race and ethnicity. And in some practical ways, Directive 15 is successful: the categories correspond to generally identifiable categories for the vast majority of Americans.

A MIXED-RACE OR "OTHER" CATEGORY

The practical considerations of adding a multirace category or some other option for people of multiple ancestry are many. First, there is the issue of the status of multiracial peoples in terms of current civil rights legislation. Given that the current approach includes only one majority group, "white," a key question would be whether multiracial people would be a protected category, with the same legal rights to representation as current minority categories, or whether multiracial responses would be coded back to one of the existing five categories.

Second, if a variety of responses are going to be reallocated to the existing or any other set of categories, any coding algorithm will present potential controversies. For example, one option that has been proposed is to redistribute multiracial people to the current single-race categories: that is, if there were a population of 80 whites, 10 blacks, and 10 multiracial individuals, the recoding would be 89 whites and 11 blacks. This algorithm does not change the relative sizes of the single-race categories, but it is an arbitrary allocation of multiracial people to single-race categories. If the 10 multiracial people each had one black parent and one white parent, the recoding algorithm totally misrepresents the multiracial people. And if a person's parents are multiracial, the complexity of classification will further increase.

Another option for a multirace category is to allow respondents to write in multiple races, but the large number of races that is likely to result only raises more questions of classification. The reallocation of a multiple-race response to a single-race category would also raise difficulties, since the respondent had clearly indicated a preference for an identity other than one of a single race (or ethnicity). If self-identification is taken as a basic principle, there are no grounds for recoding a multirace person to a single race. It is difficult to imagine any logical recoding algorithm for people who decline to provide a single-race affiliation.

In a 1993 set of congressional hearings by the House Subcommittee on

Census, Statistics and Postal Personnel on the federal racial and ethnic standards, representatives of some civil rights and ethnic organizations said that they would be open to the testing of a multirace category or the self-reporting of multiple race ancestries (Fletcher, 1994; Der, 1994). But there was opposition from some American Indian groups, who argued that inclusion of a "multirace" category would compromise the current usefulness of data on the American Indian and Alaskan Native population. According to workshop participants, American Indian groups have not been asked about their reaction to collecting data with multiple responses. Such an approach would allow separate counts of the groups, although in a more complicated manner.

Third, there are questions about the statistical reliability of any new category (see below). Small categories are more vulnerable to inaccuracies from sampling error than are large categories. They are also more vulnerable to the mismatches of numerators and denominators in rate calculations. Another factor that is related to statistical reliability concerns the variability of categorization over time. People who identify themselves as multiracial at one stage in life may identify themselves in a single race or ethnic category at another time.[4]

Fourth, there is likely to be an element of confusion and nonresponse to a mixed-race category. The majority of Americans are probably of mixed ancestry, if one defines ancestry with sufficient narrowness, so the question of how far back in time to consider would be raised. Presumably, the purpose of a mixed-race category would be to include only those people whose parents are in separate race and ethnicity categories under the definitions in the current federal standard. It would be important and perhaps difficult to make this clear to respondents.

The Census Bureau's experience with the "other" race category in the last two censuses provides some information about what the addition of such a category might entail. A large proportion (41 percent) of the write-in responses were reclassified into one of the Directive 15 categories. As noted above, for the Modified Age-Race-Sex (MARS) file that the Census Bureau prepares for the use of other agencies and researchers, people in the "other race" category are assigned the same race as another nearby (in terms of processing) person who gave the same answer to the Hispanic-origin question. There is no way to evaluate how this reclassification corresponds to people's self-perceptions or, in fact, to any other basis for classification.

Additional experience with write-in responses comes from those to the open-ended ancestry questions in the 1980 and 1990 censuses. Such questions have the advantage of flexibility for respondents and offer a means of studying trends and

[4]Of course, this possibility also exists for multiracial people under the current standards as well, but it is likely that the creation of an intermediate category would increase transitions between categories, since the "distance" between the categories would presumably be closer (the distance between black and multirace in contrast to the distance between black and white).

characteristics of self-identity. But they also have several disadvantages. First, although there are certainly some people who prefer a write-in response, there are probably also some people for whom an open-ended question may cause confusion about what they are "supposed" to respond. Thus, it is not entirely clear that an open-ended question results in higher response rates or more accurate responses.

Second, the processing costs of open-ended questions are higher than those for fixed categories. Third, if the write-in responses are to be reallocated to major groups, classification algorithms must be written, tested, and harmonized across agencies. Without reallocation to a limited group of categories, agencies face the burden of how to present hundreds of different responses, making it virtually impossible to cross-classify race with other variables. Even some apparently simple nomenclature or classification issues may be troublesome: for example, should "blacks" and "African Americans" be grouped together, and if so, under what label? Lastly, as noted above, it is not clear how existing civil rights law, written to accommodate major groups, would accommodate the varied and subtle distinctions that would result from open-ended questions. More broadly, it is also not clear what significance to attach to such distinctions: are they linguistic or cultural and do they have social consequences?

Open-ended questions may prove more feasible for major statistical agencies with large data processing resources, like the Census Bureau, than for agencies for which the collection of racial and ethnic data is only a small portion of their administrative mandate, such as the Federal Reserve Board or the Department of Veterans Affairs. Overall, however, closed-ended questions are preferable. Most surveys and administrative forms have limited space, and answer categories need to be obvious to respondents.

MEASUREMENT ISSUES

Measurement issues are at the heart of the reasons for considering revisions to Directive 15; they are also central concerns for the evaluation of possible changes. This section describes five major measurement issues: relationship between statistical systems, validation of race and ethnicity data, correspondence of self-identification and observer identification, coding, and sample size requirements for surveys.

Census data and data from other federal and nonfederal statistical systems are closely related and interdependent. For example, census data are the denominator for virtually all birth, mortality, and morbidity rates. In turn, these rates are used to develop the Census Bureau's population estimates and projections. Most federal agencies use census data as the baseline for designing sampling frames, as well as the basis for personnel pools for equal employment opportunity compliance.

However, the census is not the only source of racial and ethnic data. Other