

Center for Studies in Demography and Ecology



The Meaning and Measurement of Race in the U.S. Census: Glimpses into the Future

by

Charles Hirschman
University of Washington

Richard Alba
State University of New York at Albany

Reynolds Farley
Russell Sage Foundation

The Meaning and Measurement of Race in the U.S. Census: Glimpses into the Future

Charles Hirschman
Center for Studies in Demography and Ecology
Department of Sociology, Box 353340
University of Washington
Seattle, Washington 98195-3340

Richard Alba
Department of Sociology
State University of New York at Albany
Albany, NY 12222

Reynolds Farley
Russell Sage Foundation
112 East 64th Street
New York, NY 10021

September 16, 1999 revision

Revision of a paper presented by the first author at the annual meetings of the American Sociological Association, August 20-24, 1998, San Francisco. The final version of this paper was completed while the first two authors were Visiting Scholars at the Russell Sage Foundation. The authors thank Jason Thomas for his assistance and Carolyn Liebler for her comments on an earlier draft.

The Meaning and Measurement of Race and Ethnicity in the U.S. Census: Glimpses into the Future

INTRODUCTION

There are few, if any, methodological issues in the social sciences with more consequences for contemporary American society than the measurement of racial and ethnic categories. The procedures used to construct census or other “official” categories, the dilemmas of individuals who must choose among multiple identities, and the reciprocal influences between ethnic consciousness and census classifications are increasingly studied and debated (Kibria 1998, Nagel 1994, Portes and MacLeod 1996). Yet for two hundred years, race has been a standard item in the decennial censuses of the United States, and there did not appear to be a major problem in the measurement of race—either for those in charge of collecting government statistics or for most of the general public.

The absence of popular debate about the measurement of race for much of American history does not mean that there were no problems in assigning persons to specific race categories. It is very likely that some individuals felt that the race classification did not acknowledge their true identities, especially if they were of mixed ancestry or if their appearance did not resemble stereotypical expectations (Forbes 1990). Indeed, a very large number of people probably “passed” from one racial identity to another as a means of individual social mobility. Nonetheless, the system worked in the sense that the categories used in the census “race” question were in accord with popular perceptions, and there were only minor challenges to the measurement of race in the census.

In spite of this seeming stability, there has been a sea change in the conceptualization and meaning of race, and alongside that concept has arisen a new one, that of ethnicity (Glazer and Moynihan, 1975). A hundred years ago, most people, including a significant share of intellectuals, thought of racial classifications as biological groups akin to the species (Banton 1998, Gould 1981). Over the course of the twentieth century, there was a gradual shift in both the scientific and popular understandings of race, and at the present time, the prevailing view is that races are socially defined categories (Omi and Winant, 1994;

Hollinger, 1995), making them akin to ethnic groups. In a 1987 decision granting an Iraqi the right to sue under provisions of the Civil Rights Act of 1965, the Supreme Court ruled that ethnic groups could be considered races because of the historical uses of these terms (St. Francis College v. Al-Khazraji). According to a common definition, members of a race share some common features of physical appearance, while members of an ethnic group share cultural ones (van den Berghe, 1967). But, in either case, their status in society is determined more by how they are treated by other groups than by their common descent.

Paralleling these changes in meaning have come significant changes in the measurement of race over the last century. The once-standard practice was for individuals to be assigned into racial categories by census enumerators, who probably did not ask the census question about race since the relevant characteristics were thought to be readily observed. The major turning point came with the 1960 census when most households were asked to fill in their own census form (Taeuber and Hansen 1966). The new method of data collection meant that every person (or the household respondent who filled in the census form) could classify themselves as they wished. In spite of this dramatic change in gathering racial information, the practical implications were minor. There was an increase in persons who reported themselves as American Indians (Eschbach 1993, 1995), and more persons checked the “other race” category because they thought their identity was not included in the list of races on the census questionnaire (Harrison and Bennett 1995). But the overall patterns of racial composition suggest that most Americans identified themselves by race pretty much as census enumerators had classified them.

Another change, one that in this case has had profound implications for Americans’ understanding of their society, is the inclusion, starting in 1970, of a distinct Hispanic identity question. Hispanic origins have been conceptualized as an ethnic origin independent of a person’s racial classification. But administrative actions and popular understanding have created a social position of Hispanics almost equivalent to that of one of the major racial categories. This had led to a five-category racial and ethnic scheme that has been used over and over again to describe American society--viz., non-Hispanic whites,

non-Hispanic blacks, Hispanics, non-Hispanic Asians, and non-Hispanic Indians (pithily characterized by the historian David Hollinger [1995] as the ethnoracial “pentagon”). This classification was formally created in 1977 in “Statistical Directive Number 15 (Office of Management and Budget 1997a: Appendix 1). Officially the Hispanic origin and race are independent classifications, but the popular assumption is that virtually everyone can be fit into one, and only one, of these categories.

Perhaps the greatest changes ever with regard to the government’s classification of individuals by race will occur in the 2000 Census. Individuals will be allowed to report more than one race. This change in procedure is occurring in a very different social and political climate from that which prevailed in the past. Indeed, this change is a direct response to the challenge from individuals and groups who want to see the census better reflect the growing multiracial portion of the population (Edmonston et al. 1996, Statistics Canada and U.S. Bureau of the Census 1993, Wright 1994). This looming change has aroused concern about the impact this new inquiry will have on the racial composition of the country. Some minority group leaders feel that the new methods may lead to fewer persons identifying themselves as members of their group. Statisticians and public officials fear that continuity in one of the most important series of census data will be disrupted and that data mandated for important public purposes, such as the electoral redistricting, will shift in unexpected ways

To evaluate the likely impact of this change in the measurement of race/ethnicity and other alternatives, the Bureau of the Census has undertaken a series of methodological studies, the most important of which was the “1996 Race and Ethnic Targeted Test” (RAETT) (U. S. Bureau of the Census 1997). In this article, we appraise the varied approaches to the measurement of race and ethnicity evaluated in that test and use them to reflect on the implications of the past measurement of race and ethnicity and on those of the new one.

THE MEASUREMENT OF RACE AND ETHNICITY IN HISTORICAL PERSPECTIVE

The social-science conceptions of race and ethnicity that undergird contemporary census measurements have moved an enormous distance from the popular conceptions that motivated the original inclusion of race (Cornell and Hartmann 1998). At the time of the first census of the United States in 1790, race or color was assumed to be part of the natural order with differential entitlement for citizenship and legal standing. The first naturalization law, in 1790, stipulated that only those who were “free and white” could become citizens (Heer, 1996), and the Constitution required that enslaved persons be counted at 60 percent for the purpose of determining population and electoral apportionment. Indians who did not pay taxes were excluded but taxed Indians were counted.

Towards the end of the nineteenth century, “race” became a defining category in Western thought with the pseudo-scientific Social Darwinist philosophy (Harris 1968). Race was then equated with biologically based divisions among humankind, which, in a common view shared by many early twentieth century social scientists, circumscribed the fundamental capacities such as intelligence of the members of the different races. This led to attempts to gather more detailed racial data in the late nineteenth century. Color was added as a census classification in 1850, with the categories of white, black, and mulatto, and in 1870, Chinese and (American) Indian were added as additional categories. By 1890, the search for racial, i.e., biological, precision led to census categories on degrees of African ancestry (mulatto, quadroon, and octoroon) and in 1900 on the percentage of “white blood” (0, $\frac{1}{2}$, $\frac{1}{4}$, or $\frac{1}{8}$) for each enumerated Indian (the full text of census questionnaires are reproduced in U.S. Bureau of the Census 1973; for a review of the evolution of American censuses, see Anderson 1990).

In the 1930 Census, Mexicans were included as a category in the race classification. The stigmatizing impact of being listed as a nonwhite group in the census (the other groups were Negro, Indian, Chinese, Japanese, Filipino, Hindu, and Korean) was well understood by the Mexican American community, and after protests, the Census Bureau agreed to publish no data about the Mexican race. Over the second half of the 20th century, the view

of race as a social construction that lacks a universal, inherent meaning gained wide acceptance (Omi and Winant, 1986; Nagel, 1994). The recognition that race is not a “natural” category whose meaning can be taken for granted has opened up the issue of how best to measure race so as to capture its contemporary social significance. That issue has also gained salience from the growth of interracial marriage, especially after the Supreme Court invalidated the last state anti-miscegenation laws in 1967 (Kalmijn, 1993, Qian 1997, Sandefur 1986, Shinagawa and Pang 1996). The increasing number of mixed-race Americans has blurred what have been viewed as clear-cut boundaries separating major racial groups (Root 1992, 1996).

Also contributing to questions about the measurement of race and ethnicity is the politicization of the subject. Since census data are often used to measure social and economic problems, and to indicate the size of a political constituency, the more people counted in a group, the stronger the group’s claim on governmental attention, electoral districts and resources. The efforts to include a separate census question on Hispanic identity, beginning in the 1970 Census and the separate listing of numerous Asian nationality groups on the race question were the direct products of political mobilization (Choldin 1986, Espiritu 1992, chapter 5; Lott 1998, U.S. Bureau of the Census 1990).

In the 1980 census, a new question on ancestry or ethnic origins was added to supplement the data on race and Hispanic identity and to supplant the parental place-of-birth question that had been asked since 1870. Although the richness of new data on race, Hispanic origin and ancestry might have led to a greater understanding of the racial and ethnic roots of the American population, it has instead created a crisis of sorts. Methodological studies of the reliability of the measurement of race and ethnicity were not encouraging (U.S. Bureau of the Census 1974). Pre-tests and investigations of reliability found ambiguity and uncertainty of measurement (and meaning) of race and ethnic classifications, problems that were disguised with limited data were now painfully evident (Alba, 1990; Farley 1991, Lieberman and Waters 1993, Lieberman and Santi 1985, Levin and Farley 1982).

Because of increasing intermarriage across race and ethnic divisions, the assumption of race and ethnicity as ascribed at birth is no longer tenable. The ethnic identity of the offspring of these unions do not fit neatly into the standard census categories (Xie and Goyette 1997). Comparison of persons along the three dimensions of race, Hispanic origin, and ancestry shows that there are a significant number of black Hispanics, persons who are white (by race) and American Indian (by ancestry), and other persons of blended and mixed ethnic origins. Although the numbers are not so large as to change the findings based on the conventional ethnic comparisons, and while *ad hoc* rules can be devised to handle inconsistent cases (del Pinal 1992), there is a major crack in the conceptual framework of mutually exclusive and exhaustive race and ethnic categories.

The problem is evident in the contentious task of revising of Statistical Directive 15—the Office of Management and Budget set of rules that define race and ethnic categories for data collection and presentation by government agencies (Edmonston et al. 1996, Lott 1998, Office of Management of Budget 1997a). Some groups do not find the primary source of their group identity in any of the census questions (e.g., religious groups). A growing number of persons do not answer the questions or write in that they are just “American.” With classifications that reflect a variety of criteria (physical appearance, language, treaty status, national or regional ancestry) determined solely by individual subjective choice, it is not surprising that many find the questions difficult to answer and government agencies have difficulty explaining the logic behind them (Perlmann 1997). The arrival of immigrants with mixed backgrounds such as dark-skinned Dominicans, Spanish-speaking Filipinos, Chinese from Thailand and Indians from Guatemala further challenges the traditional classification system.

After deliberate study and efforts to bring stakeholders into the process, the Office of Management and Budget revised Statistical Directive No. 15 (Office of Management and Budget 1997b). The major change is the designation of four major race categories (American Indian or Alaskan Native, Black or African American, Native Hawaiian or Other Pacific Islander, and White) and two ethnic categories (Hispanic or Latino, and Not

Hispanic or Latino). In the 2000 Census, respondents will be able to check that they are members of one or more racial categories (Appendix A contains the proposed versions of the Hispanic origin and race questions for the 2000 Census; the most current guidelines for the measurement are reported in Office of Management and Budget 1999). Opening up the race question to multiple responses—that is, adding what is known as the Tiger Woods questions to the census questionnaire—is a dramatic new change in government data collection.

THE 1996 RACE AND ETHNIC TARGETED TEST

As part of its planning for the 2000 Census and in response to the revisions in Statistical Directive No. 15, the Bureau of the Census conducted several inquiries to determine how the American population would respond to alternative questions regarding the measurement of race and ethnic origin. The most comprehensive of these preliminary studies was the 1996 Race and Ethnic Targeted Test (RAETT) (U.S. Bureau of the Census 1997).

The 1996 RAETT used an experimental design to test the impact of eight possible changes in questionnaire format on race and ethnic responses (U.S. Bureau of the Census 1997: Appendix A). In each version, the first four items in the questionnaire were identical—the standard census questions on name, sex, date of birth and, and household relationship. But the next two questions on race and Hispanic/Latino/Spanish origin varied across the versions of the questionnaire. The experiment included variations in the order of the questions, question wording, the inclusion of a multiracial/biracial category, directions that allowed for multiple response to the race question, and a combined race and Spanish origin question. Although individuals answered only one version of the questionnaire, it is possible to measure the impact of different questionnaire formats (within the range of sampling error) because the different versions of the questionnaire were randomly assigned to households within each targeted sample. In the analysis reported below, we concentrate on the four versions of the questionnaire that are most significant: the 1990 Census standard, the 1990 Census standard with an additional category of multiracial/biracial, the

2000 Census format with the option marking more than one race, and a combined race/Hispanic origin question (the questionnaires for these four versions are reproduced in Appendix B). The RAETT survey did not obtain information about a respondent's place of birth, educational attainment nor economic status so the relationship of those variables to the reporting of race cannot be studied.

Because some race and ethnic groups comprise only a small fraction of the total U.S. population, a national probability survey would not have included enough cases with sufficient numbers to test all versions of the questionnaire for every group. Instead the Bureau designed a "targeted sample" design based on geographic concentrations for six populations. The six targeted populations were White ethnic, Black, Hispanic, American Indian, Asian or Pacific Islander, and Alaska Native (U.S. Bureau of the Census 1997: Appendix Table B-1). All but the small Alaskan Native group—only 77,000 thousand Aleuts and Eskimos were reported in the Census of 1990—will be considered here. The targeted samples for each group were drawn from a variety of geographical areas with ethnic group concentrations. Since the samples are geographically defined, they do not consist only of the targeted group. The samples were large enough, however, that there is a sufficient number of observations so that variations in response to questionnaire format can be examined.

These targeted samples are not necessarily representative of the national population for each group. Because the RAETT samples were targeted to areas of group concentration, persons who live in areas with a lower density of their race/ethnic community are likely to be under-represented. Another potentially important limitation of the RAETT is non-response to the mail questionnaire. Of the 112,100 questionnaires mailed to selected households, only 58,917 (52.6%) were returned (U.S. Bureau of the Census 1997: 3.5). The mail response rates varied from 71 percent of the White ethnic sample to 44 percent of the Hispanic sample and 34 percent of the Alaska Native sample. Hispanic households received both English and Spanish language versions of the questionnaires. In spite of the limitations of the sample and the low response rates, the RAETT data offer an important

opportunity—and the only opportunity prior to the census—to gauge how persons in different groups respond to variations in census questions on race and ethnicity. Caution dictates, however, against an excessively fine reading of small differences. The results presented here can be interpreted as general orders of magnitudes, but not exact quantities.

A FOCUSED EXAMINATION OF THE RAETT RESPONSES

In this analysis, we focus on variations in four alternative questionnaire formats—or experiments—on the measurement of race and Hispanic origin. The exact questionnaires of four panels (A, B, C and F) are reproduced in Appendix B and are represented across the columns in each of the five major tables in this article. The first column (panel A in the RAETT) is very similar to the standard race question as asked in the 1990 Census in which respondents were directed to select only one racial category. The one difference between the RAETT “standard” and the 1990 Census is that the Hispanic origin question preceded the race question in the RAETT as it will be in the 2000 census. In the 1990 Census, the race question was before the Hispanic origin question.

The second column (panel B in the RAETT) is similar to the standard one, but with an additional category, “Multiracial/Biracial.” The third column (panel C in the RAETT) is another variant that is very close to the procedure that will be used in the 2000 Census. The format and categories are the same, but the instructions ask respondents to mark “one or more” race categories. The fourth or final column (panel F in the RAETT) is a novel experiment that combines race and Hispanic/Latino origin into a single question (rather than two separate questions) and also includes the option that allows respondents to mark one or more boxes representing racial or ethnic origins. This change is confounded, however, with an abbreviated listing of only the six major racial groups (plus a “other race” category) printed on the form. Respondents were asked to write in their specific ancestry or ethnic group.

The results of these four experiments in question format are contrasted across the four columns in Tables 1 through 5. The targeted White ethnic sample is presented in Table 1,

the targeted Black sample in Table 2, the targeted American Indian sample in Table 3, the targeted Asian and Pacific Islander sample in Table 4, and the targeted Hispanic sample in Table 5. The targeted samples represent populations in areas where the targeted population is over-represented relative to other areas.

Despite the superficially complex appearance of the tables, conclusions emerge with unmistakable consistency from the results produced by the different experiments. To start with, the reporting of race by non-Hispanics appears to be affected only slightly by changes in format. The smallest changes are registered in the white and black samples. Although the estimates from the RAETT samples can not be generalized to the national population, the percentage of persons choosing multiple racial identities in the 2000 census is likely to be in the neighborhood or one to two percent in these groups.

TABLE 1 ABOUT HERE

To illustrate the general patterns, consider the results in Table 1 for the targeted white sample (the population in census tracts that were predominantly white). Note that, in this table as in the others, the responses to the race question, which total 100 percent for each question format (down each column), are organized into two broad categories—those who marked only one of the standard groups and those who marked multiracial or multiple race categories. In their responses to the standard race question in the first column, the targeted white sample reported that almost 96 percent are solely white. Reading across the columns of Table 1 shows that the alternative questionnaire formats elicited almost no variations in response from this predominately white population. The proportion selecting only the “white” response ranged from 95 to 96 percent. This sort of modest variation in responses is certainly within the range of that would be expected from random sampling within the same population.

Nor does the reporting of multiple race differ substantially across questionnaire format. Although there were not supposed to be multiple race responses to the standard question in

column 1, a little over one percent of this sample ignored the instructions to “mark only one” race and marked at least two racial categories (see the entry “unrequested more than 1”). Neither the multiracial category, given as an option in column 2, nor the “mark 1 or more races” instruction in column 3 register much divergence from the traditional question for these white respondents. Indeed, the experimental question in column 3 that explicitly invited multiple responses obtains only slightly more of them (1.4 percent) than did the standard format (1.1 percent). Although these numbers are very small in relative terms, the absolute number (two to three million persons) with multiple race identities is not a trivial figure and so, one might argue, the nation’s statistical system should allow such reporting.

TABLE 2 ABOUT HERE

A similar consistency emerges from the black sample (Table 2), which is somewhat less concentrated residentially than the targeted white ethnic sample. Based on the standard census race question in column 1, almost 72 percent of the targeted sample identify themselves as black or African American. There is no suggestion that any of the alternative question formats affect the proportion of the targeted sample that identifies as “single race” black/African American. The percent black actually rises a tiny bit in the samples with alternative question formats, but these variations are within the range of sampling errors. The RAETT test sampled African-Americans who lived in areas that had a high density of blacks. It is possible that blacks who live in largely white or Latino neighborhoods in 2000 will report multiple races more frequently than those in the RAETT. The increase in the multiple race reporting in columns 3 and 4 (and Hispanic in column 4) appear to be persons who reported themselves to be white or “other race” in the standard race question.

There is not as much consistency for the American Indian and Asian and Pacific Islander (API) samples. But the RAETT results indicate that the numbers identifying solely as American Indians or as Asian and Pacific Islanders will not be substantially reduced with the option of choosing multiple races. Indeed, both groups will gain because some persons who previously to be members of “other races” in the standard format will identify

themselves by checking American Indian (or API) **and** another group with the “mark 1 or more” format. This is a second major conclusion.

TABLE 3 ABOUT HERE

In the targeted American Indian sample, which is found in rather heterogeneous areas, only about one third of the population is Indian, almost 5 percent are members of “other races,” and more than 15 percent are Hispanic (see column 1 of Table 3). More than the targeted white and black samples, the targeted American Indian sample does produce a modest response to the multiracial category in column 2. About 3.7 percent of the sample selected the multiracial category (in column 2). That corresponds to a drop in the single American Indian category of about 3.1 percentage points from column 1 to column 2. The inclusion of the “mark 1 or more” instruction in columns 3 and 4, however, increases the number of persons with American Indian ancestry. Of the 4 percent who do check more than one race in column 3, about 3 percent include American Indian as one of the selections. With this number added to the 37 percent of the sample that select only American Indian race in column 3, then about 40 percent of the targeted sample have some American Indian identity. This is about 4 percentage points higher than the standard question (column 1) elicits, though this distinction may not be statistically significant (U.S. Bureau of the Census 1997: Table 2-3R). The potential for increased reporting of American Indian as one of several races in 2000 seems great. In 1990, two million persons reported that they were American Indian by race, but nearly nine million persons reported an American Indian ancestry or ethnic origin (Harrison and Bennett 1995: 209)

The choice of Asian and Pacific Islander (API) origins is complicated because of the many groups and the alternative procedures for measurement. The standard race question as used in the 1990 Census listed the nine specific national groups listed in column 1 of Table 4 as well as residual options: “Other Asian and Pacific Islander” and “Some Other Race.” The targeted sample included 65 percent API of which 5 percent was in the Other API category.

TABLE 4 ABOUT HERE

The inclusion of the multiracial category in column 2 does not affect any of the API groups with one exception: Hawaiians. The proportion reporting sole race “Hawaiian” drops by almost half from column 1 to column 2. This lower figure for Hawaiian is maintained with the alternative question formats (in columns 3 and 4) when marking more than one race is allowed.

There is an additional difference in the question formats between columns 3 and 4 for the targeted API sample. Column 3 includes the complete list of Asian and Pacific Islander categories and the respondent can simply mark 1 or more. In the combined race/Hispanic origin classification in column 4, there are only the five major groups listed (White, Black/African American, Indian (American)/Alaskan Native, Asian or Pacific Islander, and Hispanic/Latino/Spanish Origin) plus Some Other Race. Then respondents were asked to write in their ancestry or ethnic group (among the listed examples were Korean, Samoan, and Taiwanese).

Spokespersons for Asian groups have frequently lobbied the Census Bureau to specifically list many Asian national origin populations on the census race question. More than 20 were proposed for 1990. Listing the specific groups produces about a 5 percentage point boost in reporting an API identity (see column 3 relative to column 4). There was also an increase of about 5 to 7 percentage points in reporting solely “Other API” in columns 3 and 4 (relative to the standard question). This result is probably a result of those who checked the generic API box in column 4, but did not bother to write in a specific ethnic/ancestry group. The increase in “Other API” in column 3, which listed all of the detailed API groups, is not so easily explained.

There was a significant proportion—almost 4 percent—of this targeted sample who marked more than 1 race, even in the standard question (column 1) when not requested to do so. The number of persons who claim partial API identity was 9 percent in column 3 (10

percent – 1 percent), which contained all the detailed API categories and allowed respondents to mark more than one category. The number with partial API identity was much less in column 4, where respondents were not prompted with the list of detailed API categories to mark.

A third major finding, evident in the results discussed above, is that the multiple race option, which will be the 2000 census standard, captures the same, or even higher, numbers as the multiracial/biracial category, but with some distinct advantages. The multiple race classification allows for the identification of specific races, while the multiracial category does not. Moreover, it appears that many persons mark multiple race categories even when told to “mark only one.” The acceptance of multiple race reporting provides explicit recognition in the national statistical system for a response that is beginning to emerge spontaneously.

Finally, changes in the measurement of race/ethnicity are potentially more consequential for Hispanics than any other group. The identification of the Hispanic population exhibits a sensitivity to variations in question format and even to question order. Many Hispanics, especially immigrants, are unsure of how to respond to census questions on race because the North American concept of race is less rooted in Latin American cultures (Rodriguez 1992). Although the addition of Hispanic origin question to the 1970 Census was, in part, a response to political pressures (Choldin, 1986), the continued census measurement (and its inclusion in OMB’s Statistical Directive 15 and other statistical inquiries) has certainly increased the political consciousness of Hispanics as members of a group distinct from Whites and Black/African Americans.

TABLE 5 ABOUT HERE

The distribution of the targeted Hispanic sample by the standard race and Hispanic Origin/Ancstry is reported in column 1 of Table 5. The question format in column 1 from the 1996 RAETT is the same as those used in the 1990 census, except for a reversal in the

order of the questions. Having the Hispanic origin question before the race question means that Hispanics are less likely to be confused by a race question that has no Hispanic category. Nonetheless, about 17 percent of the targeted Hispanic sample responds with “other race” (meaning none of the above) and another 14 percent do not answer the question at all.

The impact of question order is apparent in the two RAETT versions of questions with the “multiracial category,” one with race question first (column 2a) and one with the Hispanic origin question first (question 2b). This simple difference changes the percent of the sample choosing the “other race” category by almost 10 percentage points. Without the benefit of seeing the Hispanic origin question first, almost a quarter of the sample reports being a member of a group not listed among the responses. When the Hispanic question is first, these people report themselves to be white rather than some other race (compare 2a and 2b).

This ambiguity, however, does not lead to significant numbers responding to the multiracial option (columns 2a and 2b) or marking 2 or more races (column 3). Throughout all of these options in the measurement of race, there is general stability in the responses to the standard Hispanic origin question (modest fluctuations are due normal sampling variation). About three-quarters of the targeted Hispanic sample responds with a Hispanic identity (note that the composition of specific Hispanic categories reflects the targeted sampling areas, not of the composition of the U.S. Hispanic population as a whole).

The final experiment in which Hispanic is included as a category in combined race/ethnicity question (column 4) has a dramatic impact on measurement. More than half of the sample (56 percent) selects Hispanic as their sole identity and another 19 percent select Hispanic and another group (mostly White). Non-response to the race question drops from 13-14 percent (see columns 1 to 3) to less than one percent for the combined question. If Hispanic is a choice, many persons who identify as white or black in the standard question switch to Hispanic.

One of the major reservations expressed against combining the race and Hispanic origin questions is the potential reduction in the estimation of Hispanics or of specific Hispanic groups. The standard Hispanic origin question is independent of race and lists each of the major Hispanic groups individually (Mexican, Puerto Rican, Cuban, Other Spanish/Hispanic/Latino). The conventional assumption is that this question maximizes the numbers of persons reporting Spanish or Hispanic descent. This supposition, however, is not supported by the responses to the merged race/origin question in column 4. Adding the 56.4 percent who mark Hispanic solely with the 18.8 percent (19.4 percent – 0.6 percent) who mark Hispanic *and* another race yields a total of 75.2 percent Hispanic, which is comparable to the numbers identified by the separate Hispanic origin question (74.1 to 77.5 percent).

An idiosyncratic feature of the combined question format in column 4 (panel F) was not to list the specific Hispanic origin populations. Respondents could mark the box beside Hispanic, Latino, or Spanish origin, but they had to write in specific ancestry/ethnic group in a designated space (examples included Mexican and Ecuadorian). The result was that many persons reported generic Hispanic identity, but did not write in their specific Hispanic group. Comparing the numbers who checked their specific identity for the separate Hispanic origin question (column 1) versus those who had to write it in (column 4) shows a deficit in all Hispanic groups, especially Puerto Ricans and Cubans.

It is reasonable to ask whether the reporting of Hispanic ancestry, given its sensitivity to the question stimulus, is even more affected outside of Hispanic areas, where Hispanic identity is likely not to be as salient as it is in the areas covered by the targeted Hispanic sample. Although we do not have precise data on this question, there is a hint from an earlier table that Hispanic responses may be slightly reduced in the combined question format because mixed-Hispanic individuals in non Hispanic areas might be less likely to mark multiple identities. For instance, though the number of persons in the targeted white ethnic sample who report a Hispanic origin is fairly small, the numbers bounce around from 1.2 to 1.9

percent for the separate Hispanic origin question (Table 1). When Hispanic is offered as one of the categories in a combined race/origin question, only 1 percent of the sample reported a single Hispanic identity and another 0.6 (data not shown here) of the 1.7 percent with multiple race identities include a partial Hispanic identity. In this case, the combined question counts nearly as many Hispanics as does as the separate Hispanic question. The gap is somewhat larger in the targeted black sample, where the combined question yields a total of 3.1 percent who choose Hispanic alone or Hispanic and black together, in comparison to the 4.5 to 5.0 percent who choose Hispanic with the separate Hispanic question.

Current policies call for the Census Bureau to allocate a race or Spanish-origin if the respondent leaves the question blank. In the Census of 1990, only 1.2 percent had their race allocated but 3.5 percent did not answer the Spanish-origin question so a response—typically indicating “not of Hispanic origin”—was assigned. In the formats used in 1990, both the race and Hispanic-origin questions go unanswered by non-negligible fractions of the population. In the non-Hispanic targeted samples, the race question in the standard format was not answered by 2 to 4 percent of the respondents and the Hispanic-origin question was not answered by 3 to 7 percent. In the targeted Hispanic sample, the race question is not answered by 14 to 25 percent of the respondents.

The overall amount of missing data is dramatically reduced by the combined format. This reduction occurs in all target samples. In all but the black sample, the rate of non-response falls to 1 percent or less, while it is just 1.4 percent for the black sample. Even though there are persistent concerns that Hispanic identity is not the same as a racial identity, and that separate questions are needed to maintain continuity, the decrease in non-response with the combined question suggests that this format is closer to popular understandings of “origins,” with the consequence that fewer people skip the question.

CONCLUSIONS

The shift to allow multiple race reporting in the 2000 census is another incremental step in the evolutionary process of realigning statistical categories of group identity to conform with contemporary understanding of the makeup of the population. The census, and the broader society, still uses the increasingly anachronistic “race” word, though often with quotation marks to indicate distance from the earlier meaning of biologically defined populations sharing a common descent. The contemporary interest in race, and its measurement in the census and other statistical inquiries, is largely to document socioeconomic inequalities in accord with congressional mandates for data and to provide critical information to remedy past injustices (McKenney and Cresce 1993). Measurement of intergroup inequality requires, however, consistency of measurement across data sources and over time. Herein lies the dilemma.

In theory, consistent measurement of variations in the makeup of the population requires a clear conceptualization of the differences between segments of the population and relatively simple means to observe the relevant criteria for assignment of individuals into distinct and exhaustive categories. In practice, there are problems in measuring every characteristic, including the most simple such as age, place of residence and household relationship. For most census questions, the demographic and social characteristics are considered to be objective phenomena and knowable if appropriate tools of measurement are designed. Race is different because it includes an inherently subjective component (Liebersohn 1993).

Once upon a time, race was thought to be readily observed by outward appearances. This conception died, however, with the recognition that identity and appearance were frequently not congruent. Now a person’s race is simply whatever he or she (or another household member) says it is. The shift in measurement—consistent with the changes in popular understanding—was a major step in realigning race from a biological to a social category. The impact of this change was muted, however, because most respondents continued to consider their racial identities to be an objective characteristic determined by their phenotype.

In spite of this apparent continuity, the political pressures to create new racial or quasi-racial categories have exposed the artificial nature of the concept of race and the increasingly patchwork schema of measurement (Peterson 1987, 1997). The current crisis was prompted by pressures to accommodate persons with multiple racial ancestries and others who did not see their group in the list of census categories (Edmonston et al. 1996). Although statisticians and others can argue that wholesale changes in the procedures of measuring race might wreak havoc with the continuity of prior data, this presumes that there have been objective and reliable criteria for the measurement of race in the past. There is, however, almost no claim for the validity of measuring of race beyond self-identity (the claim of present or past discriminatory treatment as a recognizable group is sometimes mentioned, but this criterion has not yet become part of the official debate). If “race,” has little or no objective basis and self identity is the only legitimate means of measurement, then it is difficult to dismiss the argument that all identities of any demographic or historic significance should be represented in a democratic society that officially recognizes “racial” groups. All of these issues came to a head with the debate over the revision of Statistical Directive 15 and the planning for the measurement of race in the 2000 Census (Office of Management and Budget 1997a, 1997b).

This analysis of RAETT data suggests that changes in the 2000 Census will not have a dramatic impact on the measured racial composition of the population, at least not in the immediate future. A small fraction of the largest groups, perhaps one or two percent of whites and blacks will identify with more than one race. Any of the proposed methods of assignment of persons with mixed ancestry by the Bureau of the Census or research analysts will result in a classification that is fairly consistent with prior censuses. Nor will there be significant changes in the numbers of American Indians or Asians who identify solely with one community. Persons with multiple ancestries may increase the numbers in these groups because some of them would have responded with another identity (white or “other race”) to the race question used in the 1990 census.

Allowing individuals to report multiple-races seems unlikely to engender the lability that is apparent in responses to the ancestry question (Hout and Goldstein 1994). Ancestry reports appear to be influenced by the specific examples listed beneath the question and by other factors (Lieberson and Waters 1993). The underlying social identities of respondents in answering the race question are “refracted” through the different question formats, but they are not refracted very much (or, at least, not yet).

Shifting the order of the race and Hispanic-origin questions to put the latter first is likely to affect the racial distribution among Hispanics but will not affect non-Hispanic groups greatly. From the RAETT test and other Census Bureau experiments, it seems highly likely that a reordering will reduce Hispanic reporting of “other” race. This is not a negligible effect, since a good deal of research reveals that the racial distinctions among Hispanics in the census have social significance (e.g., Denton and Massey, 1989).

Our analysis of the RAETT data has revealed another issue that has largely been submerged in the debate over Statistical Directive 15 and one that has not appeared in the planning for the 2000 census, namely non-response to the census race question. About 2 to 4 of the overall population does not answer this question and the number rises to 13-14 percent in predominately Hispanic areas. The most consistent finding in the RAETT study is that this number is cut to less than one percent among all segments of the population when Hispanic is included as a category in a classification that combines race and Hispanic origin. This reduction in non-response may be due primarily to changes in response among the Hispanic population, but the data are suggestive that all groups are more likely to understand the question when Hispanics are included as a category.

The inclusion of the Hispanic category into a combined classification has the advantage of changing the census question from the concept from “race”—which is almost impossible to define—to one of “origins.” Not only does “origins” seem to be closer to the popular understanding of diversity in American society (as indicated by the decline in non-response), but it might find a justification as the officially recognized set of

populations that need to be monitored to see if the effects of past official discrimination have been eliminated. Framed in this way, Hispanic identity belongs in the same classification with the other categories.

To be sure, non-response is not the only factor to consider in weighing the merits of the combined question. One cost would be a reduction in the amount of racial data for Hispanics, since, as noted, skin color correlates with important social differences among Hispanics. Another potential consequence is the reduction in the number of reported Hispanics in a combined question. Our results do not show that the impact is noticeable, if those with sole Hispanic identity and combined ethnic identity are added together. But this evidence is based primarily on the targeted Hispanic sample in predominately Hispanic areas. Our preliminary examination of the effect of question format on Hispanic identity among persons living in predominately white or black areas do not show strong effects. But the numbers were small and it is possible that persons of mixed Hispanic background for whom their Hispanic identity is not salient may be less likely to report an Hispanic identity in the combined question format. Although the numbers of Hispanics might be reduced, one might argue such results accurately reflect how individuals think of themselves. In any event, the discrepancy points up a very significant feature of the current conceptualization of the Hispanic population in census measurement. Because Hispanic origin is treated as independent of race and thus individuals are counted as Hispanic when they report Hispanic ancestry, regardless of whatever other ancestries they may possess. There is an asymmetry in the handling of Hispanic origins and other group ancestries. This asymmetry is likely to be more problematic as intermarriage and the number of individuals with mixed racial/ethnic origins grow. This issue may become important in planning for the 2010 census.

Changes in racial and ethnic reporting will, in the short run, strongly influence litigation and legislation about racial issues including affirmative action. In the longer run, the official construction of new race categories (and blended racial identities) will impact ethnic consciousness and identities in ways that cannot be imagined at the present time.

With the recognition by the governmental statistical system that multiple racial ancestry deserves to be part of the record keeping, more people may be willing to acknowledge (or seek out information about) mixed ancestry. The change in record keeping may also break down one of the last fears against inter-racial marriage—the question of where will the children fit in. In the not too distant future, social scientists may chart shifts and life course patterns in race and ethnic identity as chosen, rather than ascribed, statuses.

REFERENCES

- Alba, Richard. 1990. *Ethnic Identity: The Transformation of White America*. New Haven: Yale University Press.
- Anderson, Margo J. 1990. *The American Census : A Social History*. New Haven: Yale University Press.
- Banton, Michael. 1998. *Racial Theories*. Second edition. Cambridge: Cambridge University Press.
- Choldin, Harvey M. 1986. "Statistics and Politics: The 'Hispanic Issue' in the 1980 Census." *Demography* 23: 403-418.
- Cornell, Stephen and Douglas Hartmann. 1998. *Ethnicity and Race: Making Identities in a Changing World*. Thousand Oaks: Pine Forge Press.
- del Pinal, Jorge H. 1992. "Exploring Alternative Race-Ethnic Comparison Groups in Current Population Surveys." *Current Population Reports, Series P23-182*. Washington, D.C.:U.S. Bureau of the Census.
- Denton, Nancy and Douglas Massey. 1989. "Racial Identity Among Caribbean Hispanics: The Effect of Double Minority Status on Residential Segregation." *American Sociological Review* 54:790-808.
- Edmonston, Barry, Joshua Goldstein, and Juanita T. Lott, eds. 1996. *Spotlight on Heterogeneity: The Federal Standards for Racial and Ethnic Classification Summary of a Workshop*. Washington, D.C.: National Academy Press.
- Eschbach, Karl. 1993. "Changing Identification among American Indians and Alaskan Natives." *Demography* 30: 635-652.
- Eschbach, Karl. 1995. "The Enduring and Vanishing American Indian." *Ethnic and Racial Studies* 18: 89-108.
- Espirtu, Yen Le. 1992. *Asian American Panethnicity: Bridging Institutions and Identities*. Philadelphia: Temple University Press.
- Farley, Reynolds. 1991. "The New Census Question About Ancestry: What Did it Tell Us?" *Demography* 28:411-430.
- Forbes, J. D. 1990. "The Manipulation of Race, Caste, and Identity: Classifying AfroAmericans, Native Americans, and Red-Black Peoples." *Journal of Ethnic Studies* 17:1-51.

- Glazer, Nathan and Daniel Patrick Moynihan, eds. 1975. *Ethnicity: Theory and Experience*. Cambridge: Cambridge University Press.
- Gould, Stephen Jay. 1996. *The Mismeasure of Man*. Revised and expanded edition. New York : Norton.
- Harris, Marvin. 1968. "Race." Pp. 263-268 in *International Encyclopedia of the Social Sciences*, vol. 13, edited by David L. Sills. New York: The Macmillan Company and the Free Press.
- Harrison, Roderick and Claudette Bennett. 1995. "Racial and Ethnic Diversity" pp. 141-210 in Reynolds Farley, ed. *State of the Union: America in the 1990's Volume Two: Social Trends*. New York: Russell Sage Foundation.
- Heer, David. 1996. *Immigration in America's Future: Social Science Findings and the Policy Debate*. Boulder: Westview.
- Hollinger, David. 1995. *Postethnic America: Beyond Multiculturalism*. New York: Basic Books.
- Hout, Michael and Joshua R. Goldstein. 1994. "How 4.5 Million Irish Immigrants Became 40 Million Irish Americans: Demographic and Subjective Aspects of the Ethnic Composition of White Americans." *American Sociological Review* 59:64-82.
- Kalmijn, Matthijs. 1993. "Trends in Black/White Intermarriage." *Social Forces* 72: 119-146.
- Kibria, Nazli. 1998. "The Contested Meanings of 'Asian American': Racial Dilemmas in the Contemporary U.S." *Ethnic and Racial Studies* 21:
- Levin, Michael J. and Reynolds Farley. 1982. "Historical Comparability of Ethnic Designations in the United States." *Proceedings of the American Statistical Association 1982 Social Statistics Section*, pp. 4-14. Washington, D.C.: American Statistical Association.
- Lieberson, Stanley. 1993. "The Enumeration of Ethnic and Racial Groups in the Census: Some Devilish Principles." In Statistics Canada and U.S. Bureau of the Census. 1993. *Challenges of Measuring an Ethnic World: Science, Politics, and Reality*, pp. 23-35. Washington, D.C.: U.S. Government Printing Office.
- Lieberson, Stanley and Lawrence Santi. 1985. "The Use of Nativity Data to Estimate Ethnic Characteristics and Patterns." *Social Science Research* 14:31-46.
- Lieberson, Stanley and Mary Waters. 1988. *From Many Strands: Ethnic and Racial Groups in Contemporary America*. New York: Russell Sage.

- Lieberson, Stanley and Mary Waters. 1993. "The Ethnic Responses of Whites: What Causes Their Instability, Simplification, and Inconsistency?" *Social Forces* 72: 421-50.
- Lott, Juanita Tamayo. 1998. *Asian Americans: From Racial Categories to Multiple Identities*. Walnut Creek, CA.: Altamira Press.
- McKenney, Nampeo R. and Arthur R. Cresce. 1993. "Measurement of Ethnicity in the United States: Experiences of the U.S. Bureau of the Census." In Statistics Canada and U.S. Bureau of the Census, *Challenges of Measuring an Ethnic World: Science, Politics, and Reality*, pp. 173-221. Washington, D.C.: U.S. Government Printing Office.
- Nagel, Joanne. 1994. "Constructing Ethnicity: Creating and Recreating Ethnic Identity and Culture." *Social Problems* 41: 152-176.
- Omi, Michael and Howard Winant. 1986. *Racial Formation in the United States: From the 1960s to the 1980s*. New York: Routledge..
- Office of Management and Budget. 1997a. "Recommendations From the Interagency Committee for the Review of the Race and Ethnic Standards to the Office of Management and Budget Concerning Changes to the Standards for the Classification of Federal Data on Race and Ethnicity." *Federal Register*. Vol 62, No 131, Wednesday, July 9, 1997/Notices. pp. 36874-36946.
- Office of Management and Budget. 1997b. "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity." *Federal Register*. Vol 62, No 210, Thursday, October 30, 1997/Notices. pp. 58782-58790.
- Office of Management and Budget. 1999. "Draft Provisional Guidance on the Implementation of the 1997 Standards for the Collection of Federal Data on Race and Ethnicity (February 17, 1999)." Prepared by the Tabulation Working Group of the Interagency Committee for the Review of Standards for Data on Race and Ethnicity. Available from U.S. Bureau of the Census webpage, see: <http://www.census.gov/population/www/socdemo/race.html>
- Perlmann, Joel. 1997. *Reflecting the Changing Face of America: Multiracials, Racial Classification, and American Intermarriage*. Public Policy Brief No. 35, Annandale-on-the-Hudson, NY: The Jerome Levy Economics Institute of Bard College.
- Peterson, William. 1987. "Politics and the Measurement of Ethnicity." In *The Politics of Numbers*, pp. 187-234, edited by William Alonso and Paul Starr. New York: Russell Sage Foundation.
- Petersen, William. 1997. *Ethnicity Counts*. New Brunswick, New Jersey: Transaction Publishers.

- Portes, Alejandro and Dag McLeod. 1996. "How Shall I Call Myself? Hispanic Identity Formation in the Second Generation." *Ethnic and Racial Studies* 19: 523-547.
- Qian, Zhenchao. 1997. "Breaking the Racial Barriers: Variants in Interracial Marriage Between 1980 and 1990." *Demography* 34: 263-276.
- Rodriguez, Clara E. 1992. "Race, Culture, and Latino 'Otherness' in the 1980 Census." *Social Science Quarterly* 73:930-937.
- Root, Maria P. P., ed., 1992. *Racially Mixed Peoples in America*. Newbury Park.: Sage Publications.
- Root, Maria P. P., ed., 1996. *The Multiracial Experience: Racial Borders as the New Frontier*. Newbury Park.: Sage Publications.
- St. Francis College v. Al-Khazraji. 1987. 481 U. S., 604.
- Sandefur, Gary. 1986. "American Indian Inter-marriage." *Social Science Research* 15:347-371.
- Shinagawa, Larry Hajime and Gin Young Pang. 1996. "Asian American Panethnicity and Inter-marriage." *Ameriasia Journal* 22:127-152.
- Statistics Canada and U.S. Bureau of the Census. 1993. *Challenges of Measuring an Ethnic World: Science, Politics, and Reality*. Washington, D.C.: U.S. Government Printing Office.
- Taeuber, Conrad and Morris H. Hansen. 1966. "Self-Enumeration as a Census Method." *Demography* 3: 289-295.
- Ueda, Reed. 1994. *Postwar Immigrant America : A Social History*. New York: St. Martins Press.
- U.S. Bureau of the Census. 1973. *Population and Housing Inquiries in the U.S. Decennial Censuses, 1790-1970*. Working Paper No. 39. Washington, D.C.: Bureau of the Census.
- U.S. Bureau of the Census. 1974. *Consistency of Reporting of Ethnic Origin in the Current Population Survey*, by Charles E. Johnson, Jr. Technical Paper No. 31. Washington, D.C.: Bureau of the Census.
- U.S. Bureau of the Census. 1990. *Census of Population and Housing (1990), Content Determination Reports, 1990 CDR-6: Race and Ethnic Origin*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Bureau of the Census. 1997. *Results of the 1996 Race and Ethnic Targeted Test*. Population Division Working Paper No. 18. Washington, D.C.: Bureau of the Census.

van den Berghe, Pierre. 1967. *Race and Racism: A Comparative Perspective*. New York: Wiley.

Wright, Lawrence, 1994 "One Drop of Blood," *The New Yorker*. Vol. 70, Issue 22 (July 25): 46-50.

Xie, Yu and Kim Goyette. 1997. "The Racial Identification of Biracial Children with One Asian Parent: Evidence from the 1990 Census." *Social Forces*. 76:547-570.