

PRELIMINARY DRAFT  
Do not cite or quote  
February 1987

MEASUREMENT OF ETHNIC ORIGIN IN THE DECENNIAL CENSUS\*

by

Jeffrey S. Passel and Michael J. Levin

Population Division  
U.S. Bureau of the Census  
Washington, D.C. 20233  
(301) 763-5590

\* Paper presented at the 1987 Annual Meeting of the American Association for the Advancement of Science, session on "Whatever Happened to the Melting Pot?" Chicago, Illinois, February 14-18, 1987.

PRELIMINARY DRAFT  
Do not cite or quote  
February 1987

MEASUREMENT OF ETHNIC ORIGIN IN THE DECENNIAL CENSUS\*

by

Jeffrey S. Passel and Michael J. Levin  
Population Division  
U.S. Bureau of the Census  
Washington, D.C. 20233  
(301) 763-5590

ABSTRACT

Over the years, the Bureau of the Census has measured the ethnic origin of the population in a variety of ways. The census questions related to ethnic origin with the longest history are those on race (1790) and country of birth (1850). The newest item on ethnicity is the open-ended ancestry question that appeared for the first time in the 1980 census. Other items related to ethnicity from previous and current decennial censuses include race, Hispanic origin or descent, birthplace of parents, citizenship, year of immigration, mother tongue, and language usage. This paper reviews the history of the collection of ethnicity data in decennial censuses. We then focus on race, Hispanic origin, and ancestry data from the 1980 census to discuss the quality of the data and their comparability with the same data items from previous censuses and other data collection systems. Modifications to the race and Hispanic origin data from the 1980 census which correct for reporting errors and differences from historical definitions are described. We draw on data from the November 1979 Current Population Survey supplement on ancestry, the 1980 Content Reinterview Survey, and the 1986 National Content Test to address the accuracy and reliability of responses to the ancestry item in the 1980 census. The paper concludes with a discussion of some of the proposals for collecting ethnicity data in the 1990 census.

\* Paper presented at the 1987 Annual Meeting of the American Association for the Advancement of Science, session on "Whatever Happened to the Melting Pot?" Chicago, Illinois, February 14-18, 1987.

## MEASUREMENT OF ETHNIC ORIGIN IN THE DECENNIAL CENSUS

Jeffrey S. Passel and Michael J. Levin  
U.S. Bureau of the Census

### INTRODUCTION

Decennial censuses in the United States provide snapshots of the number, characteristics, and distribution of the population at a point in time. Not only do the results of the various censuses provide a picture of the changing American population, but the content of the census itself provides insight into salient features of American society. The choice of which questions to ask in a decennial census reflects the topics of concern to policymakers and the populace. The way in which questions are asked and data coded shows which categories are judged to be important. As the American population changes over time, the questions and responses change also. The collection of data on ethnic origin in the decennial censuses provides a prime example of this process.

Ethnic origin (or ancestry) is an extremely complicated concept to measure. It is more difficult than most other social or economic characteristics because of the lack of clear-cut definitions, changing terminology, poor reliability, and lack of knowledge on the part of respondents of the degree of affiliation with a group or groups. Some aspects of ethnicity have proved easier to measure because they are more objective and less susceptible to changes in reporting patterns; these would include place of birth, parental birthplace, and mother tongue. Until 1980, the decennial census did not include a general question on ancestry or ethnicity, but rather relied on reported racial identification or the indirect indicators mentioned above to determine ethnic origin.

In the United States, ethnicity or ancestry has a strong geographic orientation because virtually everyone in the population is descended from an immigrant. By ancestry, we generally refer to a person's ethnic group, "roots," or country in which the person, the person's parents, or ancestors were born, regardless of the number of generations removed from the country of origin. Ancestry would reflect identification, but not necessarily the degree of association or attachment with the particular ethnic group(s). As links to the country of origin become more remote in time, the degree of individual identification may lessen and make the measurement of ethnicity even more difficult. Intermarriage serves to further weaken ethnic identification because individuals end up with multiple ethnicities. As a result of factors such as these, an individual's criteria for inclusion in a particular ethnic group are often quite subjective, leading to considerable problems in the consistent measurement of ethnicity.

Decennial censuses in the United States have used four types of questions to describe the cultural and geographic origins of the population. First, data on race have been collected in every decennial census dating back to the first one in 1790. (Although the form of the question and the response categories have changed considerably.) Second, a question about birthplace has been included in the decennial censuses since 1850, allowing the determination of immigrants' countries of origin. Censuses from 1870 through 1970 also asked respondents about their parents' birthplaces, so that second-generation and third-or-higher-generation Americans could be identified. Third, censuses have included questions about language usage, such as mother tongue, ability to read English, and the use of languages other than English. Questions of this type were first asked in 1890, but

tabulations of data on foreign mother tongue date from 1910. Finally, the fourth type of question is a direct inquiry about self-identified ethnicity or ancestry. The 1970 and 1980 censuses both included a question about Hispanic (or Spanish) origin, but the 1980 census was the first to include a general, open-ended question asking respondents to report ancestry directly. Table 1 summarizes the questions related to ancestry or ethnicity in censuses since 1850.

This paper explores some of the problems of measuring ethnic origin in the decennial census. Specifically, we focus on three questions used in the 1980 census to measure ethnic origin: the race question, the Hispanic/Spanish origin question, and the ancestry question. For all three of these, we analyze data from the 1980 census and discuss plans for 1990. We treat the race and Hispanic origin questions principally in terms of reliability and historical comparability. For these two questions, the main issues relate to anomalies in the 1980 census data and how to deal with them. (See also Passel and Word, 1987 for detailed treatment of "corrections".)

The data on self-reported ancestry are treated in more detail as part of the process leading to the development of questions for the 1990 decennial census. Since the items on parental birthplace and mother tongue were eliminated from the 1980 census, analysis of the quality of self-reported ancestry has been made more difficult. In this paper, we draw on data from a November 1979 supplement on ancestry to the Current Population Survey (CPS), the 1980 census itself, the Content Reinterview Survey (CRS) for the 1980 census, and the 1986 National Content Test (NCT) to address issues relating to the quality of data on ancestry.

## RACE

### 1980 Census Data as Collected

Data on racial groups as collected in the 1980 census and published in census publications have proved difficult to use for many persons who try to combine the data with historical census data or data from other collection systems. The 1980 census results had a large number of persons (about 6.8 million) in a residual race category which was designated "Other, not specified." (See Figure 1 for a facsimile of the 1980 question and Table 2 for the data.) The vast majority of these persons were of Hispanic origin and most would have been classified as "White" in previous censuses and in many other data systems. Another large group in the residual category consisted of Asians and Pacific Islanders whose specific nationalities were not listed on the census form.

Inconsistencies between the 1980 census and previous censuses in racial definitions distort historical comparisons, particularly for Whites and for Other races. Lack of comparability between 1980 census data and vital statistics can lead to serious distortions in vital rates, especially for these same two racial groups (Chilton and Sutton, 1986). The distortions are noticeable at the national level, but can be extremely serious for local areas, particularly those with sizeable Hispanic populations.

### Preliminary OMB-Consistent Modified Race Data (1980 Census)

The Census Bureau has recognized the inconsistencies between racial classifications used in the 1980 census and other definitions. For its own postcensal population estimates and for outside data users, the Census Bureau produced "Preliminary OMB-Consistent Modified Race Data" (Passel, 1982). The

alternative race data were based on cross-tabulations by age, sex, race, and Hispanic origin from the 1980 census complete count at the state and county level. The construction of the preliminary modified race data also used information on write-in responses to the race item in 1980 census sample data at the state level. The modified race data had 6.34 million more Whites than the initial census figures, or 3.4 percent; 188,000 more Blacks, or 0.7 percent; and 6.53 million fewer persons of Other race, or a reduction of 55.9 percent. (See Tables 2 and 3.)

The "Preliminary OMB-Consistent Modified Race Data" suffer from three major limitations. First, the only tabulations available are for age, sex, race (White, Black, and Other races only), and Hispanic origin. Because they are based on aggregated data, it is difficult (or impossible) to produce consistent tabulations on characteristics other than age, sex, race, and Hispanic origin. Geographic detail is limited to the nation, states, and counties. Finally, for some of the modifications, state-specific factors were used to produce county-level data. As a result, the modifications may not be highly accurate in some counties.

#### Micro-Modified Race Data (1980 Census)

To remedy the shortcomings of the "Preliminary OMB-Consistent Modified Race Data," the Census Bureau has developed what are called "Micro-Modified Race Data" from the sample data in the 1980 census. All of the 5.95 million persons in the sample who reported in the "Other, not specified" race category were assigned to specific racial groups on the basis of other data from the census related to ethnicity, by household relationships, and eventually by a "hot deck" procedure. Virtually all of these individuals were assigned to one of the four OMB-defined race groups: White, Black,

American Indian or Alaska Native, and Asian or Pacific Islander.<sup>1/</sup> A very small number of persons (about 1,000) remained in the "other" category. (The explanation which follows is a brief summary of the work described in detail in Passel and Word, 1987.)

To make the reassignments, each person of "Other, not specified" race was subjected to a series of hierarchical edits using other census data related to racial identification:

1. Direct Assignment by RACE
2. Direct Assignment by ANCESTRY
3. Direct Assignment by MEXICAN ORIGIN
4. Assignment by Household RELATIONSHIP
5. "HOT DECK" Assignment

For direct assignment by race, certain specific write-in responses to the race question were assigned to one of the four OMB categories, for example: Laotian to Asian or Pacific Islander, Argentinian to White. These edits assigned 3.39 million persons or 56.9 percent of the original responses of "Other, not specified" race to specific racial groups. (See table 5.)

Almost all of these persons (3.18 million) were assigned as White; about 2.95 million of these persons had responded "Mexican" to the race question.

After the direct assignment by race, about 2.56 million persons remained in the "Other, not specified" category. Most of the remaining persons (70 percent) had given a Hispanic response to the race question (such as, "Puerto Rican" or "Hispano") that could not be assigned unequivocally to a racial group. Another 28 percent had responded as "Other," but had not provided a written response. The next set of edit rules made assignments to racial groups on the basis of responses to the ancestry question. For example, persons of Italian ancestry were assigned as White; persons of Afro-American ancestry, as Black; persons of Japanese ancestry, as Asian.

The ancestry set of edit rules assigned 948,000 persons, or 36.9 percent of those remaining. Most of these persons, 843,000 or 88.9 percent were assigned as White.

The third set of edit rules assigned 132,000 persons who were not yet classified but were of Mexican origin to the White category. The fourth set of edits was based on household relationships. Individuals who still remained in the "Other, not specified" category, but were related to persons with a specified race were assigned to the appropriate OMB racial groups. The rules based on household relationship assigned 325,000 persons; of these, 265,000 or 81.7 percent were assigned as White.

Even after applying the myriad of edit rules to the original 5.95 million persons in the "Other, not specified" category, a sizeable number of persons (about 1.16 million) could still not be classified. For example, many persons of Puerto Rican origin who responded as Puerto Rican to the race question could not be assigned as either White or Black by the edit rules. Another large group which remained unclassified at this point were persons who marked the "Other, not specified" category, but did not provide a written response. The remaining unclassified persons were assigned to racial groups with a "hot-decking" procedure stratified by type of Hispanic origin.

Overall, the micro-modification procedures added 5.45 million persons, or 2.9 percent to the White population in the 1980 census. (See Table 4.) The modified Black population was increased by 216,000 or 0.8 percent; the American Indian and Alaska Native population increased by 16,000 or 1.0 percent; and the Asian and Pacific Islander population by 265,000 or 7.5 percent. The principal difference, at the national level, between the

micro-modified race data and the preliminary OMB-consistent data is that the Other races category in the micro-modified data is larger by 209,000 or 3.9 percent. A major portion of this difference (about 156,000) can be traced to differences between the unmodified sample and complete counts for Asians and American Indians. (See Table 2. Passel and Berman, 1986 discuss these differences for American Indians in detail.)

One major strength of the micro-modification procedure is that there is now a file of all persons in the 1980 census sample that includes a racial designator which is consistent with the OMB definition and most other data systems. From this file, it is possible to produce tabulations for any variable and for any geographic area. Also, because the preliminary OMB-consistent procedure used state-specific factors at the county-level, corrections in the micro-modified data should be more accurate for most substate areas. However, because the file contains sample data, figures for small areas or small categories may be subject to large sampling errors. Passel and Word (1987) include a detailed description of the edit rules, tables showing the workings of the edit process, and modified racial data for subnational areas.

#### Prospects for 1990

For 1990, two versions of the race question are being considered. (See Figure 1.) Both versions are designed to collect data on all Asians and Pacific Islanders in the complete count, thus eliminating the initial shortfall in this category in the 100-percent data. The word "Race" has been reinstated as part of the question in both versions. These two forms of the question are being tested to see if they induce more persons of Hispanic origin to respond with a specific race. (See McKenney, 1986 for a discussion of the 1990 testing program for racial and ethnic items.)

In any case, it is virtually certain that a large portion of the Hispanic population will continue to respond in the Other race category in 1990. As a consequence, the race data from the 1990 census will again be inconsistent with data from censuses before 1980 and with data from many other systems.<sup>2/</sup> Thus, to produce consistent and meaningful results by race when comparing 1990 census figures with historical data series and other data systems, it will probably be necessary to modify the 1990 census data in a manner similar to what has been done for 1980. This situation is almost certainly unavoidable when data are collected by self-enumeration and self-identification in a society such as the United States' where racial and ethnic identities are not fixed and where the concepts of race and ethnicity mean different things to different people.

### HISPANIC ORIGIN

#### 1970 Census Data and Other Early Attempts

The Census Bureau first collected data on the Spanish origin population on a national scale in the November 1969 Current Population Survey. This was followed by the 1970 census which counted 9.1 million persons of Spanish origin on the basis of a self-identification question included in the 5-percent sample. (See Figure 2.) Various evaluations of the 1970 data showed the Spanish origin results to be rather poor. Only 75 percent of persons reporting Spanish origin in the 1970 CRS had answered affirmatively to the census question. In addition, various other response problems occurred, including massive misreporting into the Spanish origin category, as well as apparently serious nonresponse and erroneous omissions from the Spanish origin population. (The material in this section is drawn principally from Siegel and Passel, 1979.)

The "Central and South American origin" category in the 1970 census had a population of 1.51 million persons. By way of comparison, the November 1969 CPS had 556,000 persons of Central and South American origin; the March 1971 CPS had 501,000; and the March 1972 CPS, 599,000. (See Table 6.) These data suggest that the 1970 census had almost 1 million too many persons in this category. The geographic distribution of the Central and South American origin population showed large concentrations in the central and southern United States. Many people apparently interpreted "Central and South American" origin to mean being from the central or southern United States. Taken together with the results from the 1970 CRS, the 1970 census and the CPS results point to an extremely large number of erroneous responses on the part of non-Hispanic persons.<sup>3/</sup>

Other analyses of 1970 data suggest, however, that the Spanish origin data seriously understate the number of Hispanics in the census. Several different "Hispanic" populations were constructed from 1970 census data with other questions that could be used to ascertain ethnicity. These populations include: Spanish language, Spanish heritage, Spanish language or surname, Spanish birth or parentage, and Spanish surname. The largest of these was the Spanish language or surname population at 10.1 million (Table 7). In parts of the country with concentrations of persons of Hispanic origin, the Spanish language population and the Spanish language or surname population showed substantially larger numbers than the Spanish origin population.

A further example of the shortfall of the 1970 census is shown in Table 6. The Mexican origin population in the census was 4.5 million. In the November 1969 CPS, the March 1971 CPS, and the March 1972 CPS, the Mexican origin population exceeded 5 million, or about 10 percent more than the 1970 census.

In addition, the CPS estimates of the Cuban and Other Spanish populations were also significantly higher than the 1970 census. Thus, even though the Spanish origin population in the 1970 census included almost 1 million too many persons of Central or South American origin, it appears that the population was severely understated because many Hispanic persons who were counted in the 1970 census did not respond affirmatively to the Spanish origin question.

Data collected using the Spanish origin identifier during the 1970s did not appear to suffer from the same deficiencies as the 1970 census data.

Reinterview studies conducted in conjunction with test censuses, year-to-year match studies of the Current Population Survey, and the 1976 NCT all showed that the Spanish origin question produced consistent data. Roughly 90 percent or more of persons who responded as Hispanic in reinterviews had responded as Hispanic in the original interview.

The improvement and general acceptance of the Spanish origin question during the 1970s set the stage for its inclusion as a 100-percent item in the 1980 census. Furthermore, the continued improvement in data quality suggests that the 1970 census figures on the Spanish origin population should not be used in historical analyses. A better approximation from the 1970 census to Hispanic data collected after 1970 is the Spanish language identifier.

#### 1980 Census Data as Collected

Data from the 1980 census for the population of Hispanic origin have been shown to be of high quality. The 1980 CRS showed that approximately 90 percent of persons reporting Hispanic origin in the census did so in the reinterview, and vice versa (Fernandez, 1986). This level of consistency,

which is consistent with historical experience after the 1970 census, basically reflects the subjective nature of the identifier (Siegel and Passel, 1979). As in the past, the "Other Spanish" category proved to be the most inconsistent group. (See Figure 3 for a facsimile of the question.) Tienda and Ortiz (1986) analyzed Hispanic origin data from the 1980 census in conjunction with other items related to ethnicity. They also found substantial agreement among the items and concluded that the Hispanic origin question proved to be the best single item to identify the Hispanic population.

In spite of the overall high quality of the 1980 data on the Hispanic population, the 1980 census did show an unexpectedly large number of persons of Mexican origin in areas which historically have not had large Mexican populations, notably eastern and southern states. Furthermore, an unexpectedly large proportion of the Mexican origin population in many areas was Black. Early tabulations of sample data showed that a large proportion of these persons did not possess characteristics normally associated with Mexican ethnicity. A follow-up survey conducted to investigate this anomaly found that virtually all of the persons suspected of misreporting (over 200,000) did not consider themselves either Mexican or Hispanic. As with the "Central or South American origin" category in the 1970 census, respondents apparently misunderstood the intent of the Hispanic origin question and believed that they were responding as "American" when they marked the category "Mex.-Amer." (See U.S. Bureau of the Census, 1982a for the details of the study.)

Further analysis of the Hispanic data suggested that similar misreporting occurred in the same areas for the "Other Spanish origin" category. On the

other hand, tabulations of sample data from the 1980 census showed that about 85,000 people who did not report as Hispanic origin had three or more of the following characteristics which would normally be associated with being of Hispanic origin: Hispanic racial response, Spanish spoken in the home, Hispanic ancestry, born in a Hispanic country, and Spanish surname (5 southwestern states only). Although none of the analyses showed a massive amount of misreporting (either into or out of the Hispanic origin population) at the national level, they did indicate the possibility of substantial distortions at the local level.

#### Micro-Modified Hispanic Origin Data (1980 Census)

To measure the degree of potential misreporting and to correct the data, a procedure for evaluating and modifying the responses to the Hispanic origin item in the 1980 census was developed which analyzed responses to census questions on place of birth, language, race, ancestry, language spoken in the home, and residence 5 years ago. Passel and Word (1987) examined individual records from the sample detail file for all persons who responded as Hispanic origin plus any person with a Hispanic response to any of the items which could indicate ethnicity (race, ancestry, country of birth, language, or surname) as well as all other persons in their households. Every person in these households was assigned a score on a "Hispanicity" index which was based on responses to the ethnicity items, the current state of residence, and the state or country of birth. The more of the characteristics identifiable as Hispanic that are possessed by the individual, the higher the Hispanicity index. The specifics of the scoring system are shown in Table 8. (See Passel and Word, 1987 for more details and further explanation of the edit rules, which are summarized in the remainder of this section. See also Word, 1984.)

The quality of each individual's response to the Hispanic origin item was evaluated on the basis of the Hispanicity score taken alone, plus taken in the context of the scores and responses of other household members. Each individual was assigned a "modified Hispanic origin" on the basis of the original response, the ethnicity score, and the scores and responses of other household members. The basic strategy in this modification was to accept the original response unless there were very strong indications that the response was erroneous. Virtually all original responses (both Hispanic and non-Hispanic) were retained as the modified Hispanic origin category.

Individuals were retained as Hispanic origin if any response to any of the ethnicity items suggested that the person could be Hispanic.<sup>4/</sup> In addition, if any member of the household was retained as Hispanic, then all members of the household who responded originally as Hispanic were also retained as Hispanic, regardless of their ethnicity scores. Thus, the only way a person could be changed from Hispanic origin to non-Hispanic origin was if no household member had any Hispanic ethnic characteristic. At the national level, only 581,000 persons were changed from Hispanic origin to non-Hispanic origin. Overall, 96.0 percent of the Hispanic origin population was retained as Hispanic. (See Table 9.) However, for many areas within the country, the proportion retained was considerably smaller; for example, in the East South Central division, only 45.0 percent of the original Hispanic origin population was retained as Hispanic.

Persons who had not responded as Hispanic could be added to the micro-modified Hispanic population if their ethnicity score provided very strong evidence that the original non-Hispanic response was erroneous. For example, any two of the following would provide such evidence: a write-in

Hispanic race entry, Hispanic ancestry, a Hispanic country of birth, Spanish language spoken in the home (except the combination of place of birth and language). Other qualified combinations, i.e., those leading to an index score of 1.8 or more, can be derived from Table 8. For all additions to the modified Hispanic population, persons added were assigned to specific Hispanic categories.

In adding persons to the micro-modified Hispanic population, household relationships also played a major role. In households where some persons had originally responded as Hispanic (i.e., a mixed Hispanic/non-Hispanic household), very strong evidence was required to add a non-Hispanic person to the Hispanic origin population on the theory that the household members understood the intent of the question. As previously described, a score of 1.8 or more for a persons who had originally responded as non-Hispanic in a mixed household was required to effect a change into the modified Hispanic population.

In households where no one had originally responded as Hispanic, the first addition to the modified Hispanic population required the very same strong evidence as in mixed Hispanic households (i.e., a score of 1.8 or more). Other household members after the first one added could be added to the modified Hispanic population on lesser evidence (i.e., a score of 1.0 or more).

The edit procedures for adding to the Hispanic population shifted about 228,000 persons into the modified Hispanic population. This figure represents about 1.6 percent of the modified Hispanic population (Table 9). In fact, for the 9 states with the largest Hispanic populations--states which

account for over 85 percent of the Hispanic population, the proportion of persons added to the modified Hispanic population was exactly 1.6 percent. The rate was the same for the remaining 41 states and the District of Columbia.

The overall net effect of the modifications was to reduce the sample figure for the Hispanic population to 14.3 million or by 353,000, that is, 2.4 percent of the original sample population. Because the misresponse into the Hispanic origin tended to be in areas without large Hispanic populations, the modification increased the geographic concentration of the Hispanic population. Thus, the 9 states with the largest Hispanic populations have 87.1 percent of the modified Hispanic population as compared to 85.1 percent of the original sample Hispanic population in the 1980 census.

Results of the modification of each type of Hispanic origin (Table 10) are consistent with the evaluations from the 1980 CRS and earlier. The Puerto Rican and Cuban origin populations are highly consistent; the Mexican origin population, less so; and the Other Spanish population, even less. The modifications had very little effect on the Puerto Rican and Cuban origin populations as the modified population differed by less than one percent from the original figures. The rate at which persons were added to the population was roughly the same in all parts of the country. Deletions were, however, higher in parts of the country with few Puerto Ricans or Cubans.

Overall, the Mexican origin data also proved to be of high quality as the rate of addition was only 1.0 percent and of deletion, 3.1 percent. The geographic distribution of the deletions reflects the results of the earlier study (U.S. Bureau of the Census, 1982a) which found substantial misreporting

in eastern and southern states. The addition rate was again roughly the same everywhere. In the 6 states with the largest Mexican origin populations, there were moderate to small increases in the modified population as the deletions approximated the additions. In the rest of the country, about 1 in 5 of the original responses were assessed to be in error. (In some areas, such as the East South Central division, more than half of the original responses were modified.) The differential rate of deletions led to a greater geographic concentration of Hispanics in the micro-modified population. Whereas 87.5 percent of the original Mexican origin population lived in the 6 states with the largest numbers, 89.7 percent of the modified population did so.

The Other Spanish origin population has previously proved to be less consistently reported than other Hispanic origin groups. The micro-modified data show this same pattern. The addition rate for this modified group was about 3 percent in all areas--a rate which exceeds the other origin groups.<sup>5/</sup> The rate of deletion in the states with the largest numbers of Other Spanish approached 5 percent and was greater than the corresponding rate for other groups. In states with smaller numbers of Other Spanish, the deletion rate approached the 1 in 5 value found among the Mexican origin population. Although no direct studies of misreporting were done for the Other Spanish population as they were for the Mexican origin group, the high deletion rate suggests the presence of response problems like those that occurred in the original Mexican origin responses.

This modified Hispanic origin indicator is now appended to the record of every person in the sample detail file of the 1980 census.<sup>6/</sup> From the modified sample detail file, tabulations of modified Hispanic origin, as well

as modified race, for any variable of interest can be produced for any geographic area, subject to the limitations described previously for the micro-modified race variable. Passel and Word (1987) include a description of the edit rules and scoring system used in constructing the modified Hispanic origin identifier, tables showing the results of the modification process, and modified Hispanic data for subnational areas. Although a great deal of attention was directed to the modification process for the Hispanic data, it should be stressed that the unmodified data are of sufficient quality for most analyses, especially in areas with large Hispanic populations.

#### Prospects for 1990

Two slightly modified versions of the Hispanic origin question are being tested for use in the 1990 census (Figure 3). Both versions would provide more detailed information on the specific origins of the "Other Spanish" population in sample data. The short format also might help eliminate misreporting into the "Mexican-American" category because it is not listed as a precoded box. However, this version would make it more difficult to obtain 100-percent data on the Hispanic population by type. Both versions are being tested to see if they can continue to produce the high quality data on the Hispanic population available from the 1980 census (McKenney, 1986)

#### ANCESTRY

When the place of birth question was first asked in 1850, about 10 percent of the population was foreign-born. A high volume of immigration followed the Civil War, causing the proportion foreign-born to increase, peaking in the early decades of the 20th century (Table 11). Restrictive immigration laws

of the 1920s led to decreases in both the number and the proportion foreign-born after 1930. With the revision of immigration laws in the 1960s, immigration increased considerably and the origins of the immigrants shifted. Latin America and Asia supply a far higher proportion of immigrants today than 25 years ago.

The proportion of second-generation Americans, that is, native-born persons with 1 or 2 foreign-born parents, lags several decades behind the portion foreign-born. By 1920, 21 percent of the population were second-generation, but the proportion has decreased steadily since then. If the foreign-born population continues to grow more rapidly than the native, the long-term decline in second-generation Americans may be reversed (depending of course on the fertility rates of the native and foreign-born populations).

Data on foreign-born and foreign-parentage (i.e., the first and second generations) only describe immigration trends of a few decades. Any effect of immigration in the more distant past on ethnic composition is not measured by the birthplace and parental birthplace items when the immigrant ancestors are grandparents or more distant. The 1980 census differed from all other censuses since 1870 because it did not explicitly provide information about the sons and daughters of the foreign-born, but it did provide more general ancestry information. Since the ancestry question is an innovation and may be the most appropriate question to identify a large segment of the population, we will consider in detail the usefulness and reliability of the data.

Ancestry Data as Collected (1980 Census and Recent Surveys)

The ancestry question in the 1980 census was based on self-identification, was open-ended, and had no prelisted categories. (See Figure 4.) The November 1979 CPS and the 1986 NCT also collected ancestry data with essentially the same question.<sup>7/</sup> Some individuals reported a single ancestry group, others reported more than one group. All single- and double-ancestry responses were coded in each survey and the census.<sup>8/</sup> In addition, 17 triple-origin ancestries expected to be frequently reported were coded, while only the first two reported ancestries were coded for all other responses of three or more ancestries.

Differences between the census and the two surveys in the method of data collection can have great impact on the data. Self-enumeration was used in both the census and the NCT. Even though ancestry was based on self-identification in the CPS also, selection of response categories was not completely independent since enumerators were instructed to prod respondents for a specific ancestry if a religion, the category "American," or an unclassifiable response was given. The census and the NCT each provided a number of examples (different examples in each case), such as Hungarian, Irish, Italian, English, and Afro-American, to aid respondents; the listing may have influenced the reporting as well. In addition, the instructions on the NCT requested respondents to write "the group" with which they identified and probably led to increased reporting of single ancestries at the expense of multiple ancestry reporting.

Approximately the same proportions of the population reported as "American" or "United States" in each survey and the census, even though interviewers were instructed in the CPS to explain that "American" or a religious response

was not appropriate, since ancestry was to refer to the specific foreign nationality of the person or his or her ancestors. On the other hand, the percentage reporting a religion or other unclassifiable response was much less when enumerators were used to collect the data in the CPS. The proportion of persons not reporting any ancestry response remained fairly consistent in the three data collection operations. The proportion of nonresponse was slightly lower in the CPS than in the census. The proportion apparently increased to 13 percent in the NCT, but this difference may not be statistically significant because of the small number of cases in the NCT.

German was the largest ancestry group in the CPS and the NCT, but English was most frequently selected in the census (Table 13). The 1980 census questionnaire design may have contributed to the differences in ancestry reporting. The prominence of the term "English" in the census question on language (which immediately preceded the ancestry question) and the listing of "English" as the second example in the ancestry question may have influenced respondents to report a single ancestry of English. The contribution of single English ancestry is also seen in the large proportion of single ancestry in the census seen in Table 12. In the NCT, when the language question did not appear with the ancestry question, and when English was not on the list of examples, the percentage reporting single and multiple ancestry for English ancestry was similar to the CPS. The contribution of single English ancestry is also seen in the large proportion of single ancestry reporting in the census seen in Table 12. (Note that because persons who reported multiple ancestries were included in more than one group, the sum of persons reporting the ancestry groups was greater than the total; for example, a person reporting "German-English" was tabulated in both the "German and other group(s)" and "English and other group(s)" categories.)

### What Does the Ancestry Question Measure?

The ancestry question itself, a direct measure of ethnicity, would seem to be a substantial improvement over previous indirect measures (place of birth of person and parents and language) for identifying ethnic groups. First, a major advantage of the ancestry question is that it allows identification of the ethnicity of all persons, not just persons who would be classified on the basis of their own birthplace or birthplaces of their parents. Second, unlike birthplace, ethnicity provides data on the individual's own perception of ethnic identification. Third, the ancestry question is much more informative than the language inquiries.

The ancestry question differs from many other socio-demographic queries since responses may not be "correct" in the same manner as responses to questions about age or income. Although the accuracy of a reported ancestry cannot be determined exactly, some of the factors which are related to ancestry can be studied. The CPS, the census, and the NCT each gathered information to assist in this analysis. The 1979 CPS collected information on five items that could influence a person's selection of an ancestry--birthplace, mother's birthplace, father's birthplace, mother tongue, and current language. The census gathered data on race, Hispanic origin, birthplace, and language (which can be used primarily in assessing data for Asians and persons of Hispanic origin). The NCT gathered information only on own and parental birthplace<sup>9/</sup>, but the NCT provided additional information because the reinterview asked the ancestry questions again and the responses were matched in processing.

Several attempts have been made to develop correspondences between direct and indirect measures of ancestry by looking at generation since immigration.

For example, Levin and Farley (1982) developed correspondences between ancestry and countries of birth and languages for ancestries in their analysis of the November 1979 CPS. They found that in some situations, correspondences were easily obtained. For Korean ancestry, for example, Korea was the appropriate country of birth, and Korean was the language. Some ancestries were associated with two or more languages, such as both French and Dutch for persons reporting Belgian ancestry. Other ancestries such as West Indian or Slavic involve many countries of birth. In addition, it can be difficult to specify countries and languages for many Central and Eastern European ancestries because there have been numerous border changes and the same language is spoken in several countries.

The correspondences were used to determine generation of immigration. For example, an American-born person who reported as Italian and whose father was born in Italy became a second-generation Italian. Those persons whose birthplace or parents' birthplace did not agree with the reported ancestry were classified as third or later generations.

Table 15 presents selected ancestry groups (single and multiple combined) with more than 100,000 persons in the November 1979 CPS and illustrates one of the advantages of the ancestry question. The next-to-last column shows the proportion of the population that was third-and-later generation; in fact, the table is arranged in descending order of this percentage.

Information about a person's birthplace and parental birthplaces would only identify a small portion of some major ethnic groups. Less than 7 percent of the people who actually reported African, Irish, Dutch, Scottish, Welsh, English, Scandinavian, German, and French ancestries would be classified in those groups in 1979 if only birthplace questions were asked. Thus, the

ancestry question provides a more complete measurement for identifying these groups, especially for those whose forebears came to the United States in the 19th or early 20th centuries. On the other hand, this question is no more advantageous than the birthplace questions for identifying some of the groups that are recent arrivals, such as Korean, Iranian, Cuban, Vietnamese, or Colombian. In these instances, ancestry measures little more than the first and second generations.

It is important to point out that, although the ancestry question clearly has the potential for providing ethnic data for a very large proportion of the population, it is not clear how reliable these data are. Furthermore, persons reporting multiple ancestries may not be adequately considered in this analytic scheme. After all, if a person reports as German-English ancestry and has one parent born in Germany, the person has agreement at the second generation with one ancestry, but may or may not have agreement with the other.

Reporting of multiple ancestries can also make analysis and understanding of the data problematic. The proportion of third and subsequent generations who claim multiple ancestries illustrates the relative levels of intermarriage across ethnic groups and the consequent difficulties in measurement. (See the last column in Table 15.) In order to claim multiple ancestry, in theory, a person's parents or earlier ancestors must have been of different ancestries. In other words, in order to have a multiple ancestry, a marriage of two single ancestries should have occurred at some point. For some groups--Asian Indian, for example--about half of the persons born in the U.S. and whose parents were both born in the U.S. claimed multiple ancestries, that is, Asian Indian and some other ancestry. On the other hand, about

90 percent of the native-parentage Scottish persons claimed multiple ancestries, whereas only 6 percent of the Dominicans fell in this category.

#### Consistency of Ancestry Reporting

Since the 1979 CPS, the 1980 census, and the 1986 NCT are so different in coverage and methods of collection and processing, it is difficult to assess consistency in reporting. It is possible, however, to compare proportions reporting for 1979 and 1980 because the sample size for the November 1979 CPS is sufficiently large for some tentative conclusions. The 1986 NCT can be compared with reinterview data to assess consistency of response.

Table 16 compares 1980 census and the 1979 CPS estimates for the ancestry groups (single and multiple combined) with more than 100,000 persons in 1980. The hypothesis that each pair of estimates is congruent was tested; congruence occurred when a similar proportion of the total population gave the same ancestry response in the November 1979 CPS and the 1980 census. Specifically, the hypothesis that the 1979 and 1980 estimates were two samples drawn from one population with a common proportion giving a designated ancestry response was tested at the .05 level of significance. Those cases which tested as the same are marked with an asterisk in Table 16. For the unmarked countries, sampling alone cannot account for the differences between 1979 and 1980 data so the response patterns must have differed.

Of the 16 largest ancestry groups, i.e., those with over 2 million persons in the 1980 census, only one showed congruence between the 1979 CPS and the census. Overall, only 15 of the 50 ancestry groups did not differ between the 1979 CPS and the 1980 census. In addition, most of the 15 groups (for example, Greek, Filipino, Cuban, and Dominican) have had large influxes of immigrants in recent years and so have large proportions of first- and

second-generation Americans. Thus, they may be very cognizant of their ethnicity and consistent in reporting.

Among the single ancestry, a higher degree of congruence between the November 1979 CPS and the 1980 census can be found among the single ancestry groups as 24 of the 50 groups agree, including several of the largest ones. A notable group that does not show agreement is the English ancestry population where the census had almost twice the proportion of single ancestry responses as the November 1979 CPS. The multiple ancestry responses had 19 of 50 groups show congruence, but none of the largest ones. Although differences between the CPS and census might be expected because of differences in the methods of data collection, processing, etc. for a few groups, the differences are larger than one might expect. Since replicability between censuses is needed if ancestry is to be used as a reliable indicator of ethnic background, it is important to understand why the degree of agreement between the CPS and the census was so low for so many groups, including some numerically very large ones.

The ancestry group that was most affected by differences between the CPS and the census was English, which had about 10 million more persons in the census. This gain was registered with the addition of about 12 million more single ancestry responses and 2.7 million fewer multiple responses. The proportion reporting English ancestry rose from about 18 percent in November 1979 to 22 percent in the next April. Other parts of the British Isles were affected in the opposite direction. The Irish ancestry group lost about 3 million persons or 2 percentage points; the Scottish lost about 4 million; and the Welsh, 1 million. These losses almost completely offset the gain in the English group. Obviously, respondents did not respond the same when

conversing with an interviewer (in the CPS) and receiving cues and prodding as they did in the census when they filled out the questionnaire themselves. Multiple ancestry responses, particularly from other parts of Great Britain and Ireland were dropped by respondents when self-enumeration was used.

We have already noted the possibility that the placement of the language and ancestry questions on the form may have affected the English ancestry responses. The 1986 NCT was designed in part to eliminate this problem. Apparently it was "too successful," as the English ancestry group in the NCT was consistent with neither the November 1979 CPS or the 1980 census. The percentage reporting English ancestry in the NCT dropped considerably to 15 percent of all persons specifying an ancestry from 26 percent in the census and 22 percent in the 1979 CPS (Table 14). However, the proportion of the population giving a response of single English ancestry in the NCT was similar to the level in the CPS and much less than the census. Furthermore, the proportion of persons in the English ancestry group who gave a single ancestry response was similar in the NCT and the CPS at a level substantially lower than the census. Thus, it does appear that the language question affected the responses to the ancestry question. The problem can be cured by moving the ancestry question away from the position directly following the language question.

It is difficult to determine the consistency of reporting from aggregate data on ancestry, partly because individual responses cannot be verified. The 1986 NCT provides an opportunity to make some direct assessment of reporting consistency because a reinterview was conducted with part of the sample which repeated the ancestry question. Of the population who gave a specific ancestry response in both the NCT and the reinterview, about 75 percent gave

at least one ancestry response which matched. However, the proportion who gave exactly the same response in both interviews was much smaller. Thus, the only study available for addressing directly the consistency of ancestry reporting shows moderate consistency, but nowhere near the levels found for race and Hispanic origin reporting.

#### Prospects for 1990

The open-ended question on ancestry offers a number of advantages over other methods of collecting data on ethnicity. First, it provides more complete information about all individuals in the population, not just those who are foreign-born or who had foreign-born parents. The traditional birthplace questions can identify no more than a small fraction of some of the largest ethnic groups in the country--for example, those claiming German, Irish, or English ancestry. Second, the ancestry question minimizes confusion between birthplace and ancestry. An Italian origin family whose ancestors lived in Argentina for two generations before they moved to the United States would be classified as Argentinian using birthplace questions, whereas they would have the opportunity to identify themselves as Italian with the ancestry question. Third, the ancestry question is much more informative than language inquiries because many different ancestry groups share common languages.

Studies of the quality of ancestry data show some consistency over time in the degree of response to the question and in the type of response (i.e., single versus multiple responses). With the exception of the English category in the 1980 census (which was probably contaminated by the placement of the language question), the ancestry data show moderate consistency over time, particularly for persons who report a single ancestry. The popularity

of the question and need for ancestry data cannot be disputed. The 1990 census will contain a question on ancestry on the sample form; versions being tested are shown in Figure 4.

#### CONCLUSION

The concepts used to measure ethnic origin in decennial censuses have changed over time to meet the needs of a wide variety of data users. Current laws and regulations require the collection of data for racial groups (White, Black, Asian or Pacific Islander, and American Indian or Alaska Native) and the Hispanic population. In addition, numerous programs require data on ethnic groups not defined by race, Hispanic origin, or birthplace. To meet these demands, the Census Bureau is planning to include three questions in the 1990 census to collect the needed data--one on race, one on Hispanic origin, and a general one on ancestry--in addition to questions on place of birth and language usage. There is an active testing program in progress to ensure that the data collected are of high quality and meet the needs of the broadest range of data users.

NOTES

1. Unless otherwise noted, the terms "American Indian and Alaska Native," "American Indian, Eskimo, and Aleut," and simply "American Indian" are used interchangeably in this paper. Similarly, the terms "Asian and Pacific Islander" and "Asian" are used interchangeably.
2. In fact, unless the same number or proportion of Hispanics respond as Other race in 1990 as in 1980, the 1990 data will also be inconsistent with the 1980 race data.
3. This response problem persisted in tests for the 1980 census and ultimately led to the elimination of the "Central and South American origin" category from the listed responses to the Hispanic origin question. The misresponse into this category proved to be a precursor of a different, but similar, response problem in the 1980 census.
4. A positive ethnicity score was sufficient to retain an original response of any Hispanic origin. A positive score could be attained if any of the 5 ethnicity items had a "Hispanic" response regardless of the other 4 items.
5. There may have been other factors which led to a higher addition rate for the Other Spanish group. This population was the default group for additions. That is, any originally non-Hispanic person with a Hispanicity score high enough to be added to the Hispanic population and who could not be assigned to a specific Hispanic origin category (Mexican, Puerto Rican, or Cuban) was added to the Other Spanish category.
6. This micro-modified Hispanic indicator was used in the modification of the racial classification described previously.
7. For the 1980 census and the November 1979 CPS, the question was worded "What is ...'s ancestry?" That was changed to "What is ...'s ancestry or ethnic origin?" for the 1986 NCT.
8. For all three collections, persons who called themselves "hyphenated" Americans such as "Swedish-Americans" were tabulated according to the ancestry which was not American. Also, those who specified both an ancestry and a religion were tabulated by the ancestry only.
9. However, parental birthplace cannot be cross-tabulated against ancestry since the questions appeared on different forms in the 1986 NCT.

REFERENCES

- Chilton, Roland and Gordon F. Sutton. 1986. Classification by Race and Spanish Origin in the 1980 Census and Its Impact on White and Nonwhite Rates. The American Statistician, 40 (3, August): 197-201.
- Fernandez, Edward W. 1986. Content Reinterview Survey: Spanish Origin. U.S. Bureau of the Census, 1980 Census: Preliminary Evaluation Results Memorandum, No. 110, November 12.
- Levin, Michael J. and Reynolds Farley. 1982. Historical Comparability of Ethnic Designations in the United States. Proceedings of the Social Statistics Section of the American Statistical Association: 1982.
- McKenney, Nampeo R. 1986. Testing Program for Race and Ethnic Items. Paper presented to the Census Advisory Committees of the American Indian and Alaska Native, Asian and Pacific Islander, Black, and Hispanic Population, for the 1990 Census at the Joint Advisory Committee Meeting, Roslyn, Virginia, May 19-20.
- , Reynolds Farley, and Michael J. Levin. 1983. Direct and Indirect Measures of Ethnicity: How Different Definitions Affect the Size and Characteristics of Various Ethnic Groups. Proceedings of the Social Statistics Section of the American Statistical Association: 1983.
- Passel, Jeffrey S. 1982. Procedures for Producing Preliminary OMB-Consistent Modified Race Data from the 1980 Census by Age, Sex, and Hispanic Origin for States and Counties. Tape Technical Documentation: County Population by Age, Sex, Race, and Spanish (P)reliminary OMB-Consistent Modified Race. U.S. Bureau of the Census.
- and Patricia A. Berman. 1986. Quality of 1980 Census Data for American Indians. Social Biology 33 (3-4, Winter).
- and David L. Word. 1987. Problems in Analyzing Race and Hispanic Origin Data from the 1980 Census: Solutions Based on Constructing Consistent Populations from Micro-Level Data. Paper presented at the annual meeting of the Population Association of America, Chicago, Illinois, April-May.
- Siegel, Jacob S. and Jeffrey S. Passel. 1979. Coverage of the Hispanic Population of the United States in the 1970 Census: A Methodological Analysis. U.S. Bureau of the Census, Current Population Reports, Series P-23, No. 82.
- Tienda, Marta and Vilma Ortiz. 1986. "Hispanicity" and the 1980 Census. Social Science Quarterly 67 (1, March): 3-20.
- U.S. Bureau of the Census. 1973. Population and Housing Inquiries in U.S. Decennial Censuses, 1790-1970, Working Paper No. 39, Washington, D.C.

- . 1982a. Preliminary Evaluation of Responses in the Mexican Origin Category of the Spanish Origin Item. Persons of Spanish Origin by State: 1980, 1980 Census of Population. Supplementary Report PC80-S1-7: 14-17. Washington, D.C., August.
- . 1982b. Ancestry and Language in the United States: November 1979. Current Population Reports, Series P-23, No. 116. Washington, D.C., March.
- . 1983. Ancestry of the Population by State: 1980, 1980 Census of Population. Supplementary Report PC80-S1-10. Washington, D.C., April.
- Word, David L. 1984. Evaluating the Hispanic Origin Response in the 1980 Census: A Redefinition Through Indirect Measures. Paper presented at the annual meeting of the Southern Regional Demographic Group, Orlando, Florida, October.

Figure 1. Race Question from the 1980 Census and Versions Being Tested for 1990

1980 Census Version (with word "Race")

<p><b>4. Race</b></p> <p>Mark (X) ONE box for each person.</p> <p>If "Indian (Amer.)," print tribe.</p> <p>If "Other race," print race.</p> <p>Report the race the person considers himself/herself to be.</p>	<p>1 <input type="checkbox"/> White</p> <p>2 <input type="checkbox"/> Black or Negro</p> <p>3 <input type="checkbox"/> Chinese</p> <p>4 <input type="checkbox"/> Filipino</p> <p>5 <input type="checkbox"/> Japanese</p> <p>6 <input type="checkbox"/> Korean</p> <p>7 <input type="checkbox"/> Vietnamese</p> <p>8 <input type="checkbox"/> Indian (Amer.) - Print tribe →  <input type="text"/></p> <p>9 <input type="checkbox"/> Asian Indian</p> <p>10 <input type="checkbox"/> Hawaiian</p> <p>11 <input type="checkbox"/> Guamanian</p> <p>12 <input type="checkbox"/> Samoan</p> <p>13 <input type="checkbox"/> Eskimo</p> <p>14 <input type="checkbox"/> Aleut</p> <p>15 <input type="checkbox"/> Other race - Print race →  <input type="text"/></p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Modified 1980 Version

<p><b>4. Race</b></p> <p>Mark (X) ONE box for each person.</p> <p><b>AND</b></p> <p>If "Indian (Amer.)," print enrolled or principal tribe.</p> <p>If "Other Asian or Pacific Islander," print one group, for example: Korean, Asian Indian, Hmong, Laotian, Samoan, etc.</p> <p>If "Other race," print race.</p> <p>Report the race the person considers himself/herself to be.</p>	<p>1 <input type="checkbox"/> White</p> <p>2 <input type="checkbox"/> Black or Negro</p> <p>3 <input type="checkbox"/> Indian (Amer.) - Print tribe →  <input type="text"/></p> <p>4 <input type="checkbox"/> Eskimo</p> <p>5 <input type="checkbox"/> Aleut</p> <p>6 <input type="checkbox"/> Asian or Pacific Islander (Boxes 6 to 11)  <input type="text"/></p> <p>7 <input type="checkbox"/> Chinese</p> <p>8 <input type="checkbox"/> Filipino</p> <p>9 <input type="checkbox"/> Japanese</p> <p>10 <input type="checkbox"/> Vietnamese</p> <p>11 <input type="checkbox"/> Other Asian or Pacific Islander - Print one group →  <input type="text"/></p> <p>12 <input type="checkbox"/> Other race - Print race →  <input type="text"/></p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

New Short Version

<p><b>4. Race</b></p> <p>Mark (X) ONE box for each person.</p> <p>Report the race the person considers himself/herself to be.</p> <p>If "Asian or Pacific Islander," print one group. →  <input type="text"/></p> <p>If "Indian (Amer.)," print enrolled or principal tribe. →  <input type="text"/></p> <p>If "Other race," print race. →  <input type="text"/></p>	<p>1 <input type="checkbox"/> White</p> <p>2 <input type="checkbox"/> Black or Negro</p> <p>3 <input type="checkbox"/> Asian or Pacific Islander (Print one group, for example: Chinese, Asian Indian, Hawaiian, Laotian, Vietnamese, etc.) →  <input type="text"/></p> <p>4 <input type="checkbox"/> Indian (Amer.) - Print enrolled or principal tribe →  <input type="text"/></p> <p>5 <input type="checkbox"/> Eskimo</p> <p>6 <input type="checkbox"/> Aleut</p> <p>7 <input type="checkbox"/> Other race - Print race →  <input type="text"/></p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figure 2. Spanish/Hispanic Origin Question from the 1970 Census

13b. Is this person's origin or descent— (Fill one circle)

<input checked="" type="radio"/> Mexican	<input type="radio"/> Central or South American
<input type="radio"/> Puerto Rican	<input type="radio"/> Other Spanish
<input type="radio"/> Cuban	<input type="radio"/> No, none of these

Figure 3. Spanish/Hispanic Origin Question from the 1980 Census and Versions Being Tested for 1990

1980 Census

<p><b>7. Is this person of Spanish/Hispanic origin or descent?</b>  <b>Mark (X) ONE box for each person.</b></p>	<p>1 <input type="checkbox"/> No (not Spanish/Hispanic)                  2 <input type="checkbox"/> Yes, Mexican, Mex.-Am., Chicano                  3 <input type="checkbox"/> Yes, Puerto Rican                  4 <input type="checkbox"/> Yes, Cuban                  5 <input type="checkbox"/> Yes, other Spanish/Hispanic</p>
----------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Modified 1980 Version

<p><b>7. Is this person of Spanish/Hispanic origin?</b>  <b>Mark (X) ONE box for each person.</b>  <b>AND</b>  <b>If "Yes, other Spanish/Hispanic," print one group for example: Argentinian, Colombian, Costa Rican, Dominican, Spaniard, etc. →</b></p>	<p>1 <input type="checkbox"/> No (not Spanish/Hispanic)                  2 <input type="checkbox"/> Yes, Mexican, Mex.-Am., Chicano                  3 <input type="checkbox"/> Yes, Puerto Rican                  4 <input type="checkbox"/> Yes, Cuban                  5 <input type="checkbox"/> Yes, other Spanish/Hispanic - Print one group →</p> <div style="border: 1px dashed black; height: 20px; width: 100%;"></div>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

"New Short" Version

<p><b>7. Is this person of Spanish/Hispanic origin?</b>  <b>Mark (X) ONE box for each person.</b>  <b>AND</b>  <b>If "Yes, Spanish/Hispanic," print one group. →</b></p>	<p>1 <input type="checkbox"/> No (not Spanish/Hispanic)                  2 <input type="checkbox"/> Yes, Spanish/Hispanic (Print one group, for example: Mexican, Mex.-Am., Puerto Rican, Cuban, Argentinian, Dominican, Spaniard, etc.) →</p> <div style="border: 1px dashed black; height: 20px; width: 100%;"></div>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

**Figure 4. Ancestry Question from the 1980 Census and the 1986 National Content Test**

**1980 Census**

**14. What is this person's ancestry? If uncertain about how to report ancestry, see instruction guide.**

*(For example: Afro-Amer., English, French, German, Honduran, Hungarian, Irish, Italian, Jamaican, Korean, Lebanese, Mexican, Nigerian, Polish, Ukrainian, Venezuelan, etc.)*

**1986 National Content Test**

**11. What is this person's ancestry or ethnic origin? (See instruction guide for further information.)**

*(For example: German, Italian, Afro-Amer., Argentinean, Armenian, Cajun, Cape Verdean, Dominican, Egyptian, French Canadian, Irish, Jamaican, Korean, Mexican, Nigerian, Norwegian, Polish, Slovak, Thai, Ukrainian, etc.)*

Table 1. Population Items Related to Ethnicity on General Schedules of Decennial Censuses: 1850 to 1990

Demographic Characteristic	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990 <sup>1/</sup>
Race	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Hispanic Origin	---	---	---	---	---	---	---	---	---	---	---	---	Xs	X	X
Ancestry	---	---	---	---	---	---	---	---	---	---	---	---	---	Xs	Xs
Place of Birth	X <sup>2/</sup>	X <sup>2/</sup>	X	X	X	X	X	X	X	X	X	Xs <sup>3/</sup>	Xs	Xs	Xs
Place of Birth of Parents	---	---	X <sup>4/</sup>	X	X	X	X	X	X	Xs	Xs	Xs	Xs	---	---
Language	---	---	---	---	X	X <sup>5/</sup>	X	X <sup>5/</sup>	X <sup>5/</sup>	Xs	---	Xs	Xs	Xs	Xs
Year of Immigration	---	---	---	---	X	X	X	X	X	---	---	---	Xs	Xs	Xs
Citizenship	---	---	X <sup>6/</sup>	---	X	X	X	X	X	X	X	X <sup>3/</sup>	Xs	Xs	Xs

X Item included on all questionnaires.  
 Xs Item included on sample questionnaires.  
 --- Item not included.

<sup>1/</sup> Items listed for 1990 are items currently being proposed. The questions to be included will be selected in 1988.

<sup>2/</sup> Question asked of free inhabitants only.

<sup>3/</sup> In 1960, place of birth was asked generally on a sample basis, but on a 100-percent basis in New York and Puerto Rico. Citizenship was asked only in New York and Puerto Rico, where it was a 100-percent item.

<sup>4/</sup> Question was only whether parents were foreign-born.

<sup>5/</sup> Whether person could speak English. In 1900, this was the only question; in 1920 and 1930, this question was in addition to request for mother tongue; in 1980, in addition to language spoken in the home.

<sup>6/</sup> For males 21 years of age or over

Source: U.S. Bureau of the Census, 1973. Population and Housing Inquiries in U.S. Decennial Censuses, 1790-1970, Working Paper No. 39, Washington, D.C.: Table 1, pp. 5-9.

Table 2. Comparison of Preliminary DMB-Consistent Modified Race Data with Unmodified Complete Count and Unmodified Sample Count, by Race: 1980 Census

(Populations in thousands. All figures rounded independently)

Race Group	Preliminary Modified DMB-Consistent Count	Difference of Modified and Unmodified Count		Unmodified Sample Count	Difference of Complete and Sample Count	
		Complete Count	Amount		Percent	Amount
All races.....	226,546	226,546	--	226,546	--	--
White.....	194,713	188,372	6,341	189,035	663	0.4
Black.....	26,683	26,495	188	26,482	-13	--
Other races.....	5,150	11,679	-6,529	11,028	-651	-5.6
Specified.....	5,150	4,921	229	5,077	156	3.2
A. & P.I+...	(X)	3,500	(X)	3,543	42	1.2
Amer.Ind.**.	(X)	1,420	(X)	1,534	114	8.0
Not specified.	--	6,758	-6,758	5,952	-807	-11.9
Hispanic....	--	5,842	-5,842	5,508	-334	-5.7
Not Hispanic	--	917	-917	444	-473	-51.6

(X) Not available or not defined.

-- Represents zero or rounds to zero.

\* Asian and Pacific Islander.

\*\* American Indian, Eskimo, Aleut.

Source: Passel (1982).

Table 3. Preliminary OMB-Consistent Modified Race for Persons Originally of 'Other, not specified' Race in the 1980 Census Complete Count, by Hispanic Origin

(Populations in thousands. All figures rounded independently)

Origin	Preliminary OMB-Consistent Modified Race			
	All Races	White	Black	Other
<b>POPULATION</b>				
All origins.....	6,758	6,341	188	229
Hispanic.....	5,842	5,691	151	(X)
Not Hispanic.	917	650	38	229
<b>PERCENT OF ORIGIN GROUP</b>				
All origins.....	100.0	93.8	2.8	3.4
Hispanic.....	100.0	97.4	2.6	(X)
Not Hispanic.	100.0	70.9	4.1	25.0
<b>PERCENT OF MODIFIED RACE CATEGORY</b>				
All origins.....	100.0	100.0	100.0	100.0
Hispanic.....	86.4	89.7	80.1	(X)
Not Hispanic.	13.6	10.3	19.9	100.0

(X) Not applicable.

Source: Passel (1982).

Table 4. Comparison of Micro-Modified Race Data with Unmodified Sample Count and Preliminary OMB-Consistent Modified Complete Count, by Race: 1980 Census

(Populations in thousands. All figures rounded independently)

Race Group	Micro-Modified Sample Count	Unmodified Sample Count	Difference of Modified and Unmodified Sample		OMB-Consistent Modified Complete		Difference of Micro-Modified & 'OMB' Counts	
			Amount	Percent	Count	Percent	Amount	Percent
All races.....	226,546	226,546	--	--	226,546	--	--	--
White.....	194,488	189,035	5,453	2.9	194,713	-224	-0.1	-0.1
Black.....	26,699	26,482	216	0.8	26,683	15	0.1	0.1
Other races.....	5,359	11,028	-5,669	-51.4	5,150	209	3.9	3.9
Specified.....	5,358	5,077	281	5.5	5,150	208	3.9	3.9
A. & P.I+...	3,807	3,543	265	7.5	(X)	(X)	(X)	(X)
Amer.Ind.**.	1,550	1,534	16	1.0	(X)	(X)	(X)	(X)
Not specified.	1	5,952	-5,952	-100.0	--	1	100.0	100.0

(X) Not available or not defined.  
 -- Represents zero or rounds to zero.

\* Asian and Pacific Islander.  
 \*\* American Indian, Eskimo, Aleut.

Source: Passel (1982) and Passel and Word (1987).

Table 5. Micro-Modified Race for Persons Originally of 'Other, Not Specified' Race in the 1980 Census Sample Count, by Method of Reassignment

(Data based on a sample. Populations in thousands. All figures rounded independently)

Method of Reassignment	Micro-Modified Race					
	All Races	White	Black	Asian and Pac. Islander	American Indian* Other Races	
<b>POPULATION</b>						
Total.....	5,951	5,453	216	265	16	1
RACE (Direct).....	3,386	3,179	3	203	0	1
ANCESTRY (Direct).	948	843	49	49	7	(X)
MEX. ORIGIN (Dir.)	132	132	(X)	(X)	(X)	(X)
RELATIONSHIP.....	325	265	39	13	8	(X)
'HOT DECK'.....	1,161	1,034	126	1	0	(X)
<b>PERCENT BY MODIFIED RACE CATEGORY FOR EACH METHOD OF REASSIGNMENT</b>						
Total.....	100.0	91.6	3.6	4.4	0.3	---
RACE (Direct).....	100.0	93.9	0.1	6.0	---	---
ANCESTRY (Direct).	100.0	88.9	5.2	5.1	0.8	(X)
MEX. ORIGIN (Dir.)	100.0	100.0	(X)	(X)	(X)	(X)
RELATIONSHIP.....	100.0	81.7	11.9	3.9	2.5	(X)
'HOT DECK'.....	100.0	89.1	10.9	---	---	(X)
<b>PERCENT BY METHOD OF REASSIGNMENT FOR EACH MODIFIED RACE CATEGORY</b>						
Total.....	100.0	100.0	100.0	100.0	100.0	100.0
RACE (Direct).....	56.9	58.3	1.3	76.7	0.9	100.0
ANCESTRY (Direct).	15.9	15.5	22.6	18.4	46.7	(X)
MEX. ORIGIN (Dir.)	2.2	2.4	(X)	(X)	(X)	(X)
RELATIONSHIP.....	5.5	4.9	17.8	4.8	51.7	(X)
'HOT DECK'.....	19.5	19.0	58.3	0.2	0.7	(X)

(X) Not applicable.  
 --- Represents zero or rounds to zero.  
 \* Includes Eskimos and Aleuts.  
 Source: Passel and Mord (1987).

Table 6. Hispanic Origin Population, by Type of Spanish: 1970 Census and Current Population Surveys (CPS) of November 1969, March 1971, and March 1972

(Data are from samples. Populations in thousands)

Origin and Area	1970 Census	Current Population Survey			Estimate for April 1970 from CPS's**	Difference of census and estimate	
		November 1969	March 1971	March 1972		Amount	Percent
UNITED STATES							
Total Hispanic...	9,073	9,230	8,956	9,178	9,149	-76	-0.8
Mexican.....	4,532	5,073	5,023	5,254	5,083	-551	-10.8
Puerto Rican....	1,429	1,454	1,450	1,518	1,461	-32	-2.2
Cuban.....	545	565	626	629	590	-45	-7.7
Central or South American.	1,509	556	501	599	547	962	176.0
Other Spanish...	1,057	1,582	1,356	1,178	1,468	-411	-28.0
FIVE SOUTHWESTERN STATES*							
Total Hispanic...	5,008	5,508	5,345	5,430	5,454	-446	-8.2
Mexican.....	3,939	4,360	4,358	4,549	4,384	-445	-10.2
Puerto Rican....	60	61	(X)	(X)	(X)	(X)	(X)
Cuban.....	56	82	(X)	(X)	(X)	(X)	(X)
Central or South American.	317	170	(X)	(X)	(X)	(X)	(X)
Other Spanish...	635	835	(X)	(X)	(X)	(X)	(X)
Not Mexican.....	1,069	1,148	987	881	1,070	-1	-0.1
REMAINDER OF UNITED STATES							
Total Hispanic...	4,065	3,723	3,611	3,748	3,696	369	10.0
Mexican.....	594	713	665	705	699	-105	-15.1
Puerto Rican....	1,369	1,393	(X)	(X)	(X)	(X)	(X)
Cuban.....	488	483	(X)	(X)	(X)	(X)	(X)
Central or South American.	1,192	386	(X)	(X)	(X)	(X)	(X)
Other Spanish...	422	748	(X)	(X)	(X)	(X)	(X)
Not Mexican.....	3,471	3,010	2,946	3,043	2,997	474	15.8

(X) Not available.

\* Arizona, California, Colorado, New Mexico, and Texas.

\*\* Estimates derived by interpolation.

Note: CPS data pertain to the noninstitutional population only.

Sources: U.S. Bureau of the Census, Census of Population: 1970, Subject Reports, Final Report PC(2)-1C, Persons of Spanish Origin, U.S. Government Printing Office, Washington, D.C., June 1973.

U.S. Bureau of the Census, Current Population Reports, Series P-20, No. 213, 'Persons of Spanish Origin in the United States: November 1969,' U.S. Government Printing Office, Washington, D.C., February 1971.

U.S. Bureau of the Census, Current Population Reports, Series P-20, No. 250, 'Persons of Spanish Origin in the United States: March 1972 and March 1971,' U.S. Government Printing Office, Washington, D.C., April 1973.

Table 7. Hispanic Population According to Various Identifiers, for the United States and Selected Areas Within the United States: 1970 Census

(Populations in thousands)

Identifier	United States		Arizona	California	New York	Florida
			Colorado	New Jersey	Pennsylvania	Texas
Spanish origin.....	9,073	5,009	1,749	405		
Spanish surname.....	(X)	4,668	(X)	(X)		
Spanish language.....	9,589	5,663	1,873	451		
Spanish heritage.....	9,295	6,188	1,053	451		
Spanish language or surname...	10,115	6,188	1,873	451		
Spanish birth or parentage.....	5,242	2,322	1,739	337		

  

RATIO TO SPANISH ORIGIN (PER 100)	
Spanish origin.....	100.0
Spanish surname.....	(X)
Spanish language.....	105.7
Spanish heritage.....	102.4
Spanish language or surname...	111.5
Spanish birth or parentage.....	57.8

(X) Not applicable.

Note: Spanish language population includes all persons of Spanish mother tongue and all other persons in families in which the head or spouse reported Spanish as his or her mother tongue. The Spanish heritage population consists of persons of Spanish language or surname population in the 5 southwestern states, persons of Puerto Rican birth or parentage in the 3 Middle Atlantic states, and persons of Spanish language in the remaining 42 states and the District of Columbia.

Table 8. Point Values Assigned to Ethnicity Items in the 1980 Census, Used for Constructing Hispanicity Index the Micro-Modified Hispanic Population

Census Question	1.5	1.0	0.8	0.5	0.2	0.1	0.0	-0.1	-0.2
Point Values Assigned (Spanish = 0.5 to 1.5; Partly Spanish = 0.1 to 0.5; Neutral = -0.1 to 0.0; Not Spanish = -0.2)									
RACE	Hispanic "Other Race" Write-in entry	(X)	(X)	(X)	(X)	(X)			
ANCESTRY <sup>1/</sup>	Single Spanish, Dual Spanish, Sp.-Neutral, Sp.-Near Sp., Neutral-Sp., or Near Sp.-Sp.	Spanish- Not Spanish	(X)	Not Spanish- Spanish	Near Spanish, or Near Sp.-Neut.	Filipino, Guamanian	White, American Indian, or "Other Race" but no write-in	(X)	All Others
LANGUAGE	(X)	(X)	Spanish, or Unspecified in Sp. HH	(X)	(X)	(X)	Neutral, Neut.-Neut., or Not-Near Sp.	(X)	All Others
SPANISH SURNAME	(X)	(X)	Spanish in 5 South- western states	(X)	(X)	(X)	"Filipino," Chamarro, or Unspecified in non-Sp. HH	(X)	All Others
PLACE OF BIRTH -- FOREIGN-BORN	(X)	(X)	In Hispanic country, or Unspecified in Sp. HH	(X)	(X)	(X)	Nonresponse in S.W. state, or Not in one of S.W. states	(X)	Not Spanish in 5 S.W. States
PLACE OF BIRTH -- NATIVE-BORN	(X)	(X)	(X)	(X)	(X)	(X)	Philippines, or Guam	(X)	All Others
RESIDENCE OR LOCATION IN 1980	(X)	(X)	(X)	(X)	(X)	(X)	Unspecified in non-Sp. HH	(X)	All Others
							Selected States and MSAs based on type of Spanish	(X)	All Others
							Selected States and MSAs based on type of Spanish	(X)	All Others

<sup>1/</sup> "Near Spanish" ancestries include Filipino and Guamanian; "Neutral" include responses like "American." The order of response is considered.

(X) Not applicable.

Note: For more details, see Passel and Word (1987).

Table 9. Micro-Modified and Original Hispanic Origin Population, for Selected States: 1980 Census

(Data are based on a sample. Populations in thousands)

Region, Division, or State	Micro- Modified Sample Population	Original Sample Population	Difference		Additions		Deletions		
			Amount	Percent	Amount	Percent of Modified	Amount	Percent of Original	
United States, total...	14,251	14,604	-353	-2.4	228	1.6	581	4.0	
Nine Selected States.	12,409	12,434	-25	-0.2	199	1.6	224	1.8	
All Other States.....	1,841	2,169	-328	-15.1	29	1.6	357	16.5	
<b>REGIONS</b>									
Northeast less NY,NJ...	411	453	-42	-9.4	6	1.5	49	10.8	
Midwest less IL.....	563	635	-72	-11.4	6	1.1	79	12.4	
South less FL,TX.....	436	628	-191	-30.5	10	2.2	201	32.0	
West less CO,NM,AZ,CA	431	453	-22	-4.9	7	1.6	29	6.4	
<b>DIVISIONS</b>									
New England.....	274	299	-26	-8.7	5	1.7	31	10.2	
Mid. Atl. less NY,NJ..	137	154	-16	-10.7	2	1.3	18	11.8	
E. N. Central less IL.	377	427	-50	-11.8	4	1.0	54	12.7	
W. N. Central.....	187	208	-22	-10.4	3	1.4	24	11.7	
S. Atlantic less FL...	235	334	-99	-29.6	5	2.3	104	31.2	
E. S. Central.....	55	119	-65	-54.2	1	1.8	66	55.0	
W. S. Central less TX.	146	175	-28	-16.1	3	2.2	31	18.0	
Mount. less CO,NM,AZ..	182	185	-4	-1.9	3	1.7	7	3.6	
Pacific less CA.....	249	268	-19	-6.9	4	1.5	22	8.4	
<b>SELECTED STATES</b>									
New York.....	1,653	1,661	-8	-0.5	32	1.9	40	2.4	
New Jersey.....	486	494	-9	-1.7	8	1.6	16	3.3	
Illinois.....	617	635	-18	-2.8	5	0.9	23	3.7	
Florida.....	851	858	-7	-0.8	20	2.4	27	3.2	
Texas.....	2,996	2,983	13	0.5	38	1.3	25	0.8	
Colorado.....	341	341	0	-0.1	5	1.6	6	1.7	
New Mexico.....	482	477	5	1.1	8	1.7	3	0.7	
Arizona.....	447	444	3	0.7	8	1.8	5	1.2	
California.....	4,537	4,541	-4	-0.1	74	1.6	78	1.7	

Source: Passel and Word (1987).

Table 10. Micro-Modified and Original Hispanic Origin Population by Type of Spanish Origin, for Selected States: 1980 Census

(Data are based on a sample. Populations in thousands)

Region, Division, or State	Micro-Modified Sample Population	Original Sample Population	Difference		Additions		Deletions	
			Amount	Percent	Amount	Percent of Modified	Amount	Percent of Original
<b>MEXICAN ORIGIN</b>								
United States, total..	8,495	8,679	-184	-2.1	89	1.0	273	3.1
Selected States(6)..	7,618	7,594	24	0.3	81	1.1	57	0.8
California.....	3,629	3,613	16	0.4	41	1.1	25	0.7
Texas.....	2,758	2,745	14	0.5	28	1.0	14	0.5
Illinois.....	395	405	-9	-2.3	3	0.7	12	3.0
Arizona.....	401	398	3	0.8	6	1.5	3	0.7
New Mexico.....	230	229	1	0.6	2	1.1	1	0.5
Colorado.....	204	205	0	-0.2	2	1.0	2	1.2
All Other States....	877	1,085	-208	-19.1	8	0.9	215	19.8
<b>PUERTO RICAN ORIGIN</b>								
United States, total..	1,998	2,005	-7	-0.3	17	0.8	23	1.2
Selected States(8)..	1,795	1,794	1	0.1	14	0.8	13	0.7
New York.....	981	979	2	0.2	8	0.8	5	0.5
New Jersey.....	243	242	0	0.1	2	0.8	2	0.7
Illinois.....	132	132	0	-0.2	1	0.5	1	0.7
Florida.....	96	96	0	-0.1	1	1.1	1	0.9
California.....	93	93	0	-0.5	2	1.7	2	2.2
Pennsylvania.....	89	89	0	-0.5	1	0.7	1	1.3
Connecticut.....	88	88	0	-0.1	0	0.5	1	0.6
Massachusetts.....	75	75	0	0.0	1	0.8	1	0.8
All Other States....	203	211	-8	-3.7	2	1.2	11	5.0
<b>CUBAN ORIGIN</b>								
United States, total..	812	806	6	0.8	17	2.1	11	1.3
Selected States(5)..	727	717	10	1.5	15	2.1	5	0.7
Florida.....	479	469	9	2.0	10	2.2	1	0.2
New Jersey.....	86	84	1	1.4	1	1.7	0	0.3
New York.....	79	79	0	-0.1	2	1.9	2	2.0
California.....	64	64	0	0.2	2	2.3	1	2.2
Illinois.....	20	20	0	-0.9	0	1.3	0	2.2
All Other States....	85	90	-4	-4.8	2	2.0	6	6.7
<b>OTHER SPANISH ORIGIN</b>								
United States, total..	2,945	3,114	-169	-5.4	106	3.6	274	8.8
Selected States(7)..	2,256	2,284	-28	-1.2	84	3.7	112	4.9
California.....	752	771	-20	-2.6	30	4.0	50	6.4
New York.....	562	563	0	-0.1	22	4.0	23	4.1
New Mexico.....	250	246	4	1.6	6	2.4	2	0.8
Florida.....	212	217	-5	-2.4	8	3.8	13	6.1
Texas.....	203	202	0	0.1	10	4.9	10	4.8
New Jersey.....	147	153	-6	-4.2	5	3.1	11	7.1
Colorado.....	131	131	0	0.1	3	2.5	3	2.4
All Other States....	689	830	-141	-17.0	22	3.2	163	19.6

Source: Passel and Word (1987).

Table 11. Foreign-Born and Native Population of Foreign or Mixed Parentage:  
Censuses of 1850 to 1980

Census Year	Foreign-born population				Native of foreign or mixed parents			
	Foreign-born population		Percent of total		Native of foreign or mixed parents		Percent of total	
	Number	Percent Increase	Number	Percent Increase	Number	Percent Increase	Number	Percent Increase
1980	14,080	46.4	6.2	46.4	(X)	(X)	(X)	(X)
1970	9,619	-1.2	4.7	-1.2	23,956	11.8	11.8	-1.5
1960	9,738	-6.6	5.4	-6.6	24,312	13.6	13.6	3.1
1950	10,431	-10.5	6.9	-10.5	23,589	15.6	15.6	1.9
1940	11,657	-18.4	8.8	-18.4	23,158	17.5	17.5	-10.6
1930	14,283	1.9	11.6	1.9	25,902	21.0	21.0	14.2
1920	14,020	2.9	13.2	2.9	22,686	21.4	21.4	20.0
1910	13,630	30.5	14.8	30.5	18,898	20.5	20.5	20.8
1900	10,445	12.9	13.7	12.9	15,646	20.5	20.5	36.0
1890	9,250	38.5	14.7	38.5	11,504	18.3	18.3	39.0
1880	6,680	20.0	13.3	20.0	8,275	16.5	16.5	55.4
1870	5,567	34.5	14.4	34.5	5,324	13.8	13.8	(X)
1860	4,139	84.4	13.2	84.4	(X)	(X)	(X)	(X)
1850	2,245	(X)	9.7	(X)	(X)	(X)	(X)	(X)

(X) Not available.

Note: Natives of foreign or mixed parents includes White population only for 1870 through 1950.

Table 12. Type of Ancestry Response: 1986 National Content Test, 1980 Census, and November 1979 Current Population Survey

Type of Ancestry Response	1986 NCT		1980 Census		Nov. 1979 CPS	
	Percent	Number	Percent	Number	Percent	Number
Total.....	100.0	226,546	100.0	216,613	100.0	100.0
Reported at least one specific ancestry.....	80.0	188,302	83.1	179,078	82.7	82.7
Single ancestry.....	43.2	118,565	52.3	96,496	44.5	44.5
Multiple ancestry.....	36.8	69,738	30.8	82,582	38.1	38.1
Ancestry not specified.....	20.0	38,243	16.9	37,535	17.3	17.3
American or United States...	5.7	13,299	5.9	13,592	6.3	6.3
Other.....	1.5	1,763	0.8	195	0.1	0.1
Not reported.....	12.9	23,182	10.2	23,748	11.0	11.0

Note: Other category includes responses indicating religious groups or unclassifiable responses.

Table 13. Most Common Ancestry Groups: 1980 Census and November 1979 Current Population Survey

(Populations in thousands)

Ancestry Group	1980 Census			November 1979 CPS		
	Number	Percent of Total	Rank	Number	Percent of Total	Rank
Total.....	226,546	100.00	(X)	216,613	100.00	(X)
Persons reporting*.....	188,302	83.12	(X)	179,078	82.67	(X)
SPECIFIC ANCESTRIES						
English.....	49,598	21.89	1	40,004	18.47	3
German.....	49,224	21.73	2	51,649	23.84	1
Irish.....	40,166	17.73	3	43,752	20.20	2
Afro-American, African...	21,169	9.34	4	16,193	7.48	4
French.....	12,892	5.69	5	14,047	6.48	6
Spanish/Hispanic**.....	12,257	5.41	6	12,493	5.77	7
Italian.....	12,184	5.38	7	11,751	5.42	8
Scottish.....	10,049	4.44	8	14,205	6.56	5
Polish.....	8,228	3.63	9	8,421	3.89	10
American Indian.....	6,716	2.96	10	9,900	4.57	9
Dutch.....	6,304	2.78	11	8,121	3.75	11
Swedish.....	4,345	1.92	12	4,866	2.26	12
Norwegian.....	3,454	1.52	13	4,120	1.90	13
Russian.....	2,781	1.23	14	3,466	1.60	14
Czech***.....	1,892	0.84	15	1,695	0.78	16
Hungarian.....	1,777	0.78	16	1,592	0.73	18
Welsh.....	1,665	0.73	17	2,568	1.19	15
Danish.....	1,518	0.67	18	1,672	0.77	17
Portuguese.....	1,024	0.45	19	946	0.44	..
Swiss.....	982	0.43	..	1,228	0.57	19
Austrian.....	949	0.42	..	1,070	0.49	20
French-Canadian.....	780	0.34	..	1,053	0.49	21

(X) Not applicable.

.. Not available.

\* Includes persons who reported at least one specific ancestry.

\*\* Includes all persons who reported a Hispanic ancestry or nationality. The 1980 figure may overstate slightly the number of persons reporting as Hispanic because it double counts persons who reported two specific Hispanic ancestries.

\*\*\* For 1980 census, includes persons reporting 'Czech' or 'Czechoslovakian.'

Note: Table includes all ancestry groups with at least one million persons at either date; only groups with more than one million are ranked. Sum of percentages can exceed 100 because persons reporting multiple ancestries are included in more than one group. The CPS estimates are based on population totals updated from the 1970 census and so are not consistent with the 1980 census; the CPS figures pertain to the civilian noninstitutional population only.

Sources: U.S. Bureau of the Census, Current Population Reports, Series P-23, No. 116, 'Ancestry and Language in the United States: November 1979,' March 1982.  
U.S. Bureau of the Census, 1980 Census of Population, Supplementary Report PC80-S1-10, 'Ancestry of the Population by State: 1980,' April 1983.

Table 14. Comparison of Type of Ancestry Response for Selected Ancestry Groups: November 1979 Current Population Survey (CPS), 1980 Census, and 1986 National Content Test (NCT)

Ancestry Group	Percent of all persons specifying at least one ancestry*			Percent of all persons specifying a single ancestry			Percent of all persons specifying a multiple ancestry			Percent of specific ancestry group giving single response		
	1986 NCT	1980 Census	1979 CPS	1986 NCT	1980 Census	1979 CPS	1986 NCT	1980 Census	1979 CPS	1986 NCT	1980 Census	1979 CPS
German.....	23.1	26.1	28.8	23.7	15.1	17.8	22.8	44.9	41.8	37.1	36.5	33.2
English.....	14.8	26.3	22.3	13.0	20.0	11.9	15.8	37.1	34.5	31.7	47.9	28.7
Irish.....	12.0	21.3	24.4	11.7	8.7	10.1	20.0	42.8	41.2	35.3	25.7	22.3
Italian.....	4.5	6.5	6.6	7.7	5.8	6.3	4.1	7.6	6.8	62.6	56.5	52.0
French.....	4.4	6.8	7.8	2.2	2.6	3.2	5.7	14.1	13.3	17.8	23.8	21.7
Polish.....	3.2	4.4	4.7	4.5	3.2	3.6	3.6	6.3	6.0	51.6	46.3	41.5
Scottish.....	3.1	5.3	7.9	2.0	1.0	1.7	6.5	12.7	15.2	23.1	11.7	11.4

\* Percentages may add to more than 100 percent because multiple ancestries are included in more than one group.

Table 15. Number of Persons, Proportion by Generation, and Proportion Reporting Multiple Ancestry in Third-and-Later Generations, for Selected Ancestry Groups: November 1979 Current Population Survey

Ancestry Group	Total Population	Proportion in Generation...				Percent of third or later generations reporting a multiple ancestry
		All Generations	First	Second	Third or Later	
Total*	241,302	100.0	4.4	8.9	86.7	73.5
African	2,926	100.0	0.6	0.2	99.2	9.8
Irish	43,752	100.0	0.6	3.3	96.1	79.5
Dutch	8,121	100.0	1.0	3.2	95.8	85.4
Scottish	14,205	100.0	1.1	3.4	95.5	90.4
Welsh	2,569	100.0	1.1	3.8	95.1	83.3
English	4,004	100.0	1.6	4.0	94.4	72.2
Scandinavian	340	100.0	-	6.3	93.7	70.3
German	51,649	100.0	1.5	4.9	93.6	69.1
French	14,047	100.0	1.5	5.5	93.1	80.8
Norwegian	4,120	100.0	2.1	10.1	87.8	76.4
Danish	1,673	100.0	1.8	11.0	87.3	78.8
Swedish	4,886	100.0	1.5	11.7	86.8	82.4
Swiss	1,229	100.0	3.4	10.4	86.3	81.0
Belgian	448	100.0	2.7	16.6	80.8	81.8
Finnish	616	100.0	7.3	18.0	74.8	72.4
Slavic	722	100.0	3.0	22.1	74.8	72.2
Polish	8,421	100.0	4.3	21.0	74.7	71.6
Czechoslovakian	1,695	100.0	4.8	21.6	73.5	64.3
Lithuanian	832	100.0	5.3	23.2	71.4	80.2
Ukrainian	525	100.0	9.0	20.1	70.9	71.0
Italian	11,751	100.0	6.5	25.4	68.1	66.5
Russian	3,465	100.0	6.2	27.5	66.3	70.3
French-Canadian	1,053	100.0	8.7	25.1	66.2	57.7
Hungarian	1,592	100.0	6.5	27.3	66.2	82.4
Other Spanish	3,566	100.0	19.2	15.3	65.5	49.9
Portuguese	946	100.0	13.4	21.2	65.4	68.8
Lebanese	322	100.0	9.5	27.1	63.4	64.8
Rumanian	334	100.0	11.9	30.6	57.5	76.2
Austrian	1,070	100.0	8.4	37.2	54.4	85.7
Greek	990	100.0	43.4	5.1	51.5	68.9
Yugoslavian	468	100.0	19.5	31.9	48.6	63.4
West Indian	193	100.0	28.4	27.2	44.4	47.7
Mexican	6,682	100.0	28.1	29.4	42.5	20.7
Canadian	610	100.0	21.4	39.1	39.5	78.8
Japanese	680	100.0	27.7	32.9	39.5	36.0
Chinese	704	100.0	43.7	28.3	27.9	67.0
Dominican	119	100.0	51.5	22.9	25.6	6.3
Asian Indian	182	100.0	57.5	18.5	24.1	46.5
Filipino	764	100.0	50.8	30.3	18.9	80.8
Puerto Rican	133	100.0	43.1	44.7	12.2	51.5
Jamaican	184	100.0	68.9	21.7	9.3	22.1
Colombian	118	100.0	68.1	24.6	7.3	22.1
Vietnamese	198	100.0	83.2	9.7	7.1	57.4
Cuban	676	100.0	69.4	26.3	4.4	66.8
Iranian	118	100.0	64.5	31.4	4.1	54.2
Korean	266	100.0	76.4	21.3	2.3	29.5

- Represents zero or rounds to zero.

\* Does not include American, Afro-American, or American Indian since birthplace is overwhelmingly United States. Includes persons who reported a multiple ancestry; since such persons are tabulated in more than one ancestry group, the total is more than the total population in the November 1979 Current Population Survey.

Note: The CPS estimates are based on population totals updated from the 1970 census and so are not consistent with the 1980 census; the CPS figures pertain to the civilian noninstitutional population only.

Source: McKenney, Farley, and Levin (1983).

Table 16. Proportions Reporting in Specific Ancestry Groups, by Type of Response:  
1980 Census and November 1979 Current Population Survey

(Proportions based on total population)

Ancestry Group	Total		Single ancestry		Multiple ancestry	
	1980 Census	1979 CPS	1980 Census	1979 CPS	1980 Census	1979 CPS
German.....	21.73	23.84	7.92*	7.92	13.81	15.92
Irish.....	17.73	20.20	4.56*	4.51	13.17	15.69
English.....	21.89	18.47	10.48	5.31	11.41	13.16
Scottish.....	4.44	6.55	0.52	0.76	3.92	5.81
French.....	5.69	6.48	1.35*	1.41	4.34	5.08
Afro-American.....	9.25	6.12	6.09	6.41	NA	NA
American, U.S.....	6.09	6.09	9.06	5.74	0.19	0.39
Italian.....	5.38*	5.42	3.04	2.82	2.35	2.60
American Indian.....	2.96	4.57	0.85	0.94	2.12	3.62
Polish.....	3.63	3.89	1.68	1.48	1.95	2.27
Dutch.....	2.78	3.75	0.62*	0.63	2.16	3.12
Mexican.....	3.40	3.08	3.09	2.72	0.31	0.37
Swedish.....	1.92	2.26	0.57*	0.56	1.35	1.69
Norwegian.....	1.52	1.90	0.56*	0.57	0.97	1.33
Russian.....	1.23	1.60	0.74*	0.80	0.44	0.65
Spanish.....	1.19	1.45	0.61	0.69	0.62	0.91
African.....	0.09	1.35	0.07	1.22	0.02	1.35
Welsh.....	0.73	1.19	0.14	0.21	0.60	0.98
Czech**.....	0.84*	0.78	0.14*	0.21	0.49	0.42
Danish.....	0.67	0.77	0.19*	0.20	0.48	0.57
Hungarian.....	0.78*	0.73	0.32	0.25	0.46*	0.49
Puerto Rican.....	0.64*	0.62	0.56*	0.51	0.08*	0.10
Swiss.....	0.43	0.57	0.10	0.14	0.33	0.42
Austrian.....	0.42	0.49	0.15*	0.18	0.27*	0.32
French-Canadian.....	0.34	0.49	0.15	0.27	0.15	0.22
Greek.....	0.42*	0.46	0.27*	0.26	0.15	0.20
Portuguese.....	0.45*	0.44	0.27	0.23	0.18*	0.21
Lithuanian.....	0.33	0.38	0.15*	0.15	0.18	0.24
Filipino.....	0.35*	0.35	0.28*	0.24	0.07	0.11
Slavic.....	0.08	0.33	0.03	0.14	0.04	0.19
Chinese.....	0.39	0.33	0.33	0.25	0.07*	0.08
Japanese.....	0.33*	0.31	0.29	0.24	0.05*	0.07
Cuban.....	0.26*	0.31	0.22*	0.26	0.04*	0.05
Finnish.....	0.27*	0.28	0.12*	0.12	0.15*	0.17
Canadian.....	0.20	0.28	0.10*	0.11	0.10	0.18
Ukrainian.....	0.32	0.24	0.17	0.11	0.15*	0.14
Yugoslavian.....	0.16	0.22	0.09	0.13	0.07*	0.08
Belgian.....	0.16	0.21	0.05*	0.05	0.10	0.15
Scandinavian.....	0.21	0.16	0.11	0.05	0.10*	0.11
Rumanian.....	0.32	0.15	0.17	0.06	0.15	0.09
Lebanese.....	0.13*	0.15	0.08*	0.08	0.05*	0.07
Korean.....	0.17	0.12	0.15	0.11	0.01*	0.02
Vietnamese.....	0.09*	0.09	0.09*	0.08	0.01*	0.01
West Indian.....	0.06	0.09	0.05*	0.06	0.01	0.03
Jamaican.....	0.11	0.08	0.10	0.07	0.01*	0.01
Asian Indian.....	0.14	0.08	0.12	0.07	0.01*	0.01
Dominican.....	0.08*	0.05	0.07*	0.05	0.01*	0.01
Iranian.....	0.05*	0.05	0.05*	0.05	0.01*	0.01
Columbian.....	0.07*	0.05	0.06*	0.05	0.01*	0.01
European.....	0.08	0.03	0.06	0.02	0.01	0.01

NA Not applicable.

\* Difference between 1979 and 1980 not statistically significant at .05 level.

Note: Sum of percentages can exceed 100 because multiple ancestries are included in more than one group. The CPS estimates are based on population totals updated from the 1970 census and so are not consistent with the 1980 census; the CPS figures pertain to the civilian noninstitutional population only.

Source: McKenney, Farley, and Levin (1983).