



Policy Note

2015 / 8

THE US CENSUS ASKS ABOUT RACE AND ETHNICITY: 1980–2020

JOEL PERLMANN and PATRICK NEVADA

The US Census Bureau is conducting an impressive research program that tests new formulations for questions about racial and ethnic origin. The outcome will probably be a single question that combines two current questions of great importance, on race and Hispanic origin. We applaud both the research effort and this probable outcome. Nevertheless, we call attention to the need for additional changes in ethno-racial classification, smaller and much easier to implement, that are not being discussed. Until 1980, the census included two questions that ascertained the respondent's father's and mother's country of birth; these questions should be restored to the enumeration. And the question on ancestry should be considered for elimination as redundant if the new, combined ethno-racial question is adopted.

The 1980 Census: Four Innovations

We are struck by the historical continuities between the changes now under discussion for the 2020 Census and four innovations introduced in 1980. The combined question now being tested is best seen as a fix for problematic aspects of innovations introduced in 1980. The changes of

Senior Scholar JOEL PERLMANN is director of the Institute's Immigration, Ethnicity, and Social Structure research program, and Levy Institute Research Professor at Bard College. PATRICK NEVADA is a senior at Bard College concentrating in sociology.

The Levy Economics Institute is publishing this research with the conviction that it is a constructive and positive contribution to the discussion on relevant policy issues. Neither the Institute's Board of Governors nor its advisers necessarily endorse any proposal made by the authors.

1980 were the first of two major turning points in the federal collection of ethno-racial data since the civil rights movement; the other, implemented in Census 2000, allowed individuals to declare themselves in more than one race category.

The best-known innovation of 1980 was the addition of the Hispanic Origin question. The second change was the addition of the Ancestry question. These two questions were remarkably similar in their conception of ethnic origin, but one was focused explicitly on Hispanic origin and the other covered all origins (Figure 1a).

The instructions published with the questions show just how close they were in conception:

Question 7. Hispanic Origin

A person is of Spanish/Hispanic origin or descent if the person *identifies* his or her ancestry with one of the listed groups, that is, Mexican, Puerto Rican, etc. Origin or descent (ancestry) may be viewed as the nationality group, the lineage, or country in which the person or the person’s parents or ancestors were born.

Question 14. Ancestry

Print the ancestry group with which the person *identifies*. Ancestry (or origin or descent) may be viewed as the nationality group, the lineage, or the country in which the person or the person’s parents or ancestors were born before their arrival in the United States.

The two questions ask about origin in identical terms: both concern the group with which the respondent “*identifies*”; both relate the identification to the same three terms (ancestry, origin, or descent); and both define the meaning of these three concepts using identical words (see the last sentences of the instructions for the Hispanic Origin and Ancestry questions).

The third 1980 innovation was the elimination of two other questions, no doubt partly to keep constant the total number of questions on ethno-racial origin. These asked about parental birthplace: Where was your mother born? Where was your father born? They had been on every decennial census for a hundred years.

And finally, the fourth 1980 change was that in presenting the Race question, the Bureau dropped the label “race” and only asked “Is this person . . . ?” followed by the specific categories (Figure 1b).

Figure 1a 1980 Census: New Questions on Hispanic Origin and Ancestry

7. Is this person of Spanish/Hispanic origin or descent?
 Fill one circle

- No (not Spanish or Hispanic)
- Yes, Mexican, Mexican-Amer., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, other Spanish/Hispanic

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

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

Figure 1b 1980 Census: Question on Race

4. Is this person —
 Fill one circle

- White
- Black or Negro
- Japanese
- Chinese
- Filipino
- Korean
- Vietnamese
- Indian (Amer.)
- Asian Indian
- Hawaiian
- Guamanian
- Samoan
- Eskimo
- Aleut
- Other – Specify

Print tribe

Source: www.census.gov/history/pdf/1980_long_questionnaire.pdf

The striking thing about the two new questions (Hispanic Origin and Ancestry) is that they ask about historical origins going back in time an unspecified number of generations—ignoring generational standing. An implication was that if the respondent mentioned these origins, the origins were relevant, no matter how distant in time. In this sense, they were analogized to African American origins—important, but for reasons that had nothing to do with a recent time of arrival.

By contrast, consider the two questions that were eliminated, the parental birthplace questions. These questions had carried a different implication; namely, it is important to track immigrant origins for families during the initial generations after arrival—that is, track the immigrants themselves and track separately their American-born children. Origins of later-generation descendants, by implication, are not so important

given patterns of assimilation. Moreover, given patterns of intermarriage, after several generations most people are descendants of multiple ethnic origins.

So these were the competing conceptions of the two new questions (Hispanic Origin and Ancestry) and of the two questions they replaced. Why, then, were the origin questions swapped in 1980? A crucial explanation concerns the contemporaneous thinking about Mexican American families. Recall that these changes were being urged throughout the 1970s by advocates of the groups that would soon be labeled “Hispanic.” At that time, a great many Mexican American families had lived in this country for numerous generations. Indeed, some of these families had lived in the Southwest even before the region was conquered from Mexico. These Mexican American families were concentrated in small towns and cities of the Southwest, isolated, discriminated against, and disproportionately poor. Viewed in this way, these people seemed rather more like African Americans than like descendants of Irish immigrants. But unlike African Americans, they were “invisible” to the statistical system: the census presented them as native-born whites of native parentage.

Conrad Taeuber, an important figure at the Census Bureau during those years, later recalled the challenge the Bureau faced in regard to the Mexican Americans of the Southwest. He alluded to their long history in the region, the discrimination they faced, and their distinctive socioeconomic circumstances and culture, and ruminated,

Identifying them as native born of native parentage didn't quite do it. . . . We argued that we had [could identify] native-born of foreign parentage; that gives us Mexicans of first and second generation . . . but we couldn't get away from the people in New Mexico, Arizona, and South Texas. (US Census Bureau 2003)¹

These considerations help explain the form the Hispanic Origin question took. But what explains the form and acceptance of the Ancestry question, the general question on ethnic origins? We believe there were three reasons why Ancestry joined Hispanic Origin on the 1980 questionnaire. First, Bureau officials surely felt that if self-reports on ethnic origin were a reasonable topic for the census, they should not be investigated in only one kind of ethnic group (Hispanics). Since 1969, when they began to test questions on “Spanish origin,” they also tested

general questions on “ethnic origin” or “origin or descent.” Second, there was political pressure coming from white ethnic groups that was not unlike the pressure coming from Hispanic groups. Indeed, the term “white ethnics” became very prominent after 1967. In the summary description of sociologist John Skrentny, the term applied loosely “to the mostly Catholic [or Orthodox Christian] immigrants or persons with ancestry from eastern or southern Europe . . . while Jews and Catholic Irish Americans are on the boundaries” (Skrentny 2004, 275). White Anglo-Saxon Protestants and Scandinavians were excluded. During the 1980 election year, “the ethnic desk” at President Carter’s White House apparently insisted on keeping the Ancestry question in the census. Finally, many people, especially in government, perceived a continuum of white and nonwhite minority groups. Polish and other “white ethnic” groups filed briefs on this theme in the *Regents of the University of California v. Bakke* case during the late 1970s, and indeed, the central passage in Justice Powell’s swing opinion in that case sounds rather like one of those briefs. It would have been natural for officials at the Bureau to be aware of the discussions about this continuum of groups, and how both kinds of groups were faring.

Finally, consider the fourth change made in 1980. This was the removal of the label “race” from the relevant question (from what the Bureau calls the question stem). It seems this innovation of dropping the label was introduced for narrow reasons. It was also a short-lived innovation, reversed in the next census cycle, again for narrow reasons.² Nevertheless, intended or not, the change itself—that is, dropping the classificatory label—could carry a broad message. It could suggest that the categories found under the question stem were not easily related to one covering concept. Instead, the categories were there because distinctions of different kinds had arisen in the course of American history, and a respondent’s being in one or another category affected his or her life chances.

Indeed, cutting the categories loose from any covering concept is perhaps also one useful way to think about the change that came in Census 2000, which allowed respondents to report origins in more than one race category. In any case, the acceptance of reporting multiple racial origins that year did transform the context within which a single-format question could be considered. Not least, combining the Race and Hispanic Origin questions would not involve privileging one of these two statuses over the other.

The 2020 Census: Combining Race and Hispanic Origin, Retrieving Parental Birthplace

Consider now the Bureau’s intriguing research program on the Race and Hispanic Origin questions for 2020. The single-format question being tested combines the Race and Hispanic Origin questions (Figures 2a and 2b). The new combined question would not be called the Race question but perhaps something like “Race or Origin.”³ Moreover, an alternative way to refer to the new combined question would be the way the 1980 census handled the Race question: to eliminate the classificatory label altogether. The question would then simply ask, “Which [of the following] categories describe Person 1?”

Another feature being tested for this new single-format question concerns a line under each broad ethno-racial group. The respondent is asked to write in specific origins here. Notice that what will be gathered on this line should logically duplicate the data now gathered through the Ancestry question.

The Ancestry question has had much less value than the Race and Hispanic Origin questions, for numerous reasons. First, it is not required for federal programs. Second, any unique results we get from the Ancestry question concern the ethnic identification of later-generation descendants of European immigrants and black Americans (see the examples in Figure 2a). Given high levels of intermarriage *within* both white and black America, the value of the reported ancestry data has always been questionable. But whether valuable or not, the write-in line on the 2020 question should be able to provide that information as well as the Ancestry question does. Indeed, note that data on the specific national origins of Hispanics and Asians and on the tribal affiliations of Native Americans will also come only from the write-in lines under *those* broad ethno-racial categories (Figure 2a).

We therefore anticipate that the new single-format question will replace not only the Race and Hispanic Origin questions but also the Ancestry question. Despite the minor value of the Ancestry question today and its anticipated redundancy by 2020, its supporters will probably want evidence of that redundancy before giving it up. And so the Ancestry question might be better considered for elimination around 2022 rather than before 2020. But either way, the new single-format question may yet replace three current questions—Ancestry, Race, and Hispanic Origin—rather than only the last two. In either case, the savings in respondent time and federal expenditure provide a strong rationale for bringing back the parental birthplace questions.

Figure 2a A Combined Ethno-racial Origin Question Being Tested for the 2020 Census

8. What is Person 1’s race or origin?
 Mark one or more boxes **AND** print origins.

White – Print, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc. ↴

Hispanic, Latino, or Spanish origin – Print, for example, Mexican or Mexican American, Puerto Rican, Cuban, Salvadoran, Dominican, Colombian, etc. ↴

Black or African Am. – Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc. ↴

Asian – Print, for example, Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, etc. ↴

American Indian or Alaska Native – Print name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc. ↴

Native Hawaiian or Other Pacific Islander – Print, for example, Native Hawaiian, Samoan, Chamorro, Tongan, Fijian, Marshallese, etc. ↴

Some other race or origin – Print race or origin. ↴

Figure 2b Alternative Version Omitting the “Race or Origin” Label in the Question Stem and Instructing “Mark All that Apply”

8. Which categories describe Person 1?
 Mark all boxes that apply **AND** print details in the spaces below. Note, you may report more than one group.

Source: Jones 2015

Getting these questions back is critically important now and for the foreseeable future. The most crucial point about the Hispanic origin population since the 1980s is how much it has grown through immigration. And, of course, the same can be said about the Asian American population. Today, far more Hispanics than in the 1970s are immigrants themselves or children of immigrants. Moreover, vastly more are living in places other than the old small towns and cities of the Southwest. Thus, the 1970s argument that federal statistics were missing the later-generation Mexican Americans, particularly those of the old Southwest, is less relevant to today's Hispanics. The irony here is that the questions eliminated in 1980—the parental birthplace questions—are critically relevant to the vast population of recent Hispanics and Asians. But since we do not have those questions on the relevant questionnaires, we cannot isolate the second generation.

We do *not* need these parental birthplace questions on the 100 percent census enumeration; they would have little value there. In order to understand where we do need the parental birthplace questions, the nature of three different Census Bureau enumerations must be understood. There is the decennial census, which seeks to enumerate 100 percent of the country's inhabitants on a given day. It is mandated in the Constitution and utilized for the periodic redrawing of congressional district boundaries. But this gigantic survey actually includes few questions relevant to the social and economic characteristics of the population. Consequently, a second survey, known as the American Community Survey (ACS), targets a 1 percent sample of the population each year. As such, it includes millions of people annually (and when results from several years are aggregated, many millions more). The ACS evidence is reliable down to quite small geographic areas. It covers (in addition to all the questions asked in the 100 percent enumeration) social and economic characteristics in great detail, with, for example, multiple questions on education, occupation, self-employment, unemployment, welfare, the many sources of annual income, homeownership, family and household poverty levels, housing characteristics, and specific kinds of institutionalization. The ability to discern the progress of the children of immigrants on this rich array of measures is what is at stake.

The parental birthplace questions do appear in the third relevant kind of survey, the Current Population Survey (CPS; administered by the Census Bureau for the Bureau of Labor Statistics). However, it is important to appreciate two differences

between the CPS and the ACS: size and targeted population. The CPS samples include some 60,000 households, a new set of households every 18 months. The ACS includes some 3.5 million households every year. The radically smaller size of the CPS makes it much less useful for any analysis below the national or regional level. The census data are reliable for towns, parts of cities, and rural parts of states. The size factor also makes the CPS much less useful even at the *national* level for the analysis of complex factors—for example, a study of the many factors that go into determining Mexican American poverty levels in central cities.

The second great advantage of the ACS over the CPS is the targeted population of each. Whereas the ACS counts the entire population, the CPS is restricted to the civilian noninstitutional population. When we want to know the fate of the second generation, it is important to ask, for example, how many are enlisted in the military or how many are incarcerated. Indeed, in the absence of the parental birthplace questions on the ACS, we cannot say what proportion of native-born Hispanics (or Mexican Americans) are second generation, how many of those second-generation members dropped out of high school, and how many of those dropouts are incarcerated.

These birthplace questions could be added to the ACS with little pretesting. Besides a hundred years of experience using these questions on the decennial censuses of 1880–1970, they are used regularly in the CPS. The Race and Hispanic Origin questions appear on both the decennial census and the 1 percent annual ACS. The Ancestry question appears on the ACS only. Swapping the first two of the current questions for a new single-format ethno-racial question would save much more respondent time and federal expenditure than restoring the parental birthplace questions would cost (even if the Ancestry question was not dropped).

Dropping “Race” and “Origins” from the Single-Format Question Stem

Finally, would the new single-format question be introduced with a label like “race or origin” (Figure 2a)? Or would the question stem remain unlabeled: “Which categories describe person 1?” (Figure 2b). As we noted earlier, the decision to avoid labeling the covering classification (“the question stem”) seems to have been taken in 1980 on narrow grounds and abandoned in 1990 on other narrow grounds (both irrelevant to the form of the

question in 2020). In any case, today, any number of categories, rather than one only, could be selected.

It is worth citing the way the Bureau's Nicholas A. Jones, who directs the testing program, summarized the research agenda on the labeling issue:

The 2015 NCT [National Content Test] will also evaluate the use of different conceptual terms (e.g., origin, ethnicity, or no terms) in the wording of questions. Recent Census Bureau qualitative research found that the terms "race," "ethnicity," and "origin" are confusing or misleading to many respondents, and they mean different things to different people. The 2010 AQE [Alternative Questionnaire Experiment] tested the removal of the term "race" from the question and showed no evidence that removal of the term had any effect on either unit or item response rates. Recent cognitive research tested an open-ended instruction ("Which categories describe you?") and found that respondents did not have issues with understanding what the question was asking. Therefore, an alternative option being explored tests the removal of the terms "race," "origin," and "ethnicity" from the question stem and instructions. Instead, a general approach asks, "Which categories describe Person 1?" (Jones 2015)⁴

These results suggest that nothing consequential would be lost by embracing, as was done in 1980, an unlabeled question. But note that the labels are problematic not only because respondents may be confused about them. Also at issue is the way an authoritative federal institution *uses* the terms when communicating with every American household. The "race" term has carried the implication of some sort of biological and anthropological meaning for a very long time. It is no comfort to be told that a diligent respondent can find—somewhere in the Census Bureau (and Office of Management and Budget [OMB]) texts—a disclaimer that in asking the origins question the Bureau does not mean to convey the meanings found in biological or anthropological science. What meanings for these terms *does* the Bureau (or OMB) mean to convey? And even in the new single-format question, just what is the relation between the write-ins (e.g., national origins, tribal affiliations) on the lines below the categories (e.g., White, Black, Hispanic, American Indian) and the categories themselves? If the labels

still have important conceptual references, why must races be merged with other kinds of origins in one question?

Now is the opportune moment to put all these conceptual distinctions aside in the unified question. Doing so, instead of using a question stem that includes "race or origin," will not influence the uses of the data for purposes of civil rights in general, affirmative action in particular, or any other federal program. The absence of any label may also help convey an important point. The many categories of this single question will not fit under one covering concept. Nevertheless, the categories of the unified question have all mattered in American history, and descent through one or another continues to affect social and economic well-being today.

Notes

1. A more extensive discussion of the historical evolution of the federal ethno-racial classification system, as well as a complete list of sources, can be found in Perlmann and Nevada (2015).
2. The Bureau's *Procedural History of the 1980 Census* tells us that the decision to drop the label "race" in 1980 "was made at the suggestion of the Bureau's advisory committees, which had noted that some of the categories listed in the question are not generally considered racial groups" (US Census Bureau 1986, 12-10). Presumably, the advisory committees were struck by the large number of Asian national origins as well as specific Pacific Island origins (e.g., Samoan) as races. The *Procedural History* also noted, as do many federal sources, that "the concept of race used in recent censuses reflects self-identification by respondents; it does not denote any clear-cut, scientific definition of biological stock" (US Census Bureau 1986, 12-9). In 1990, the Bureau returned to labeling the Race question as such as part of the effort "to make the intent of the question clearer and improve reporting" (McKenney and Cresce 1993, 174; US Census Bureau 1996, 14-16).
3. A critical reason for undertaking this research initiative has been the large proportion of Hispanic Origin respondents (roughly two-fifths) who have been choosing "Some other race" rather than one of the listed races since the Hispanic Origin question was introduced in 1980. Prior to that time, Mexican Americans who had chosen "Some other race" were typically reclassified by the Bureau as white; apparently,

the number involved is unknown. The choice suggests limitations in the categories; it also creates an immediate problem. Many federal agencies require age-race-sex tabulations from the census for planning purposes. These tabulations in turn recognize only the four broad race categories (Black, White, Asian, Native American) mandated by the OMB, not “Some other race.” The Census Bureau therefore creates a set of “modified” tabulations; in these, it allocates many millions of Americans (virtually all of whom are Hispanics) from “Some other race” to one of the four mandated categories.

4. According to Jones (2015), the NCT is “our primary mid-decade opportunity to compare different design options for race and ethnicity prior to making final decisions about the content of the 2020 census” and the AQE was “the most comprehensive research effort on race and Hispanic Origin ever undertaken by the Census Bureau.”

References

- Jones, N. A. 2015. “Update on the U. S. Census Bureau’s Race and Ethnic Research for the 2020 Census.” *US Census Bureau Newsletter* 3, no. 6 (April). Available at www.census.gov/people/news/issues/vol3issue6.html.
- McKenney, N. R., and A. R. Cresce. 1993. “Measurement of Ethnicity in the United States: Experiences of the U.S. Census Bureau.” In *Challenges of Measuring an Ethnic World: Science, Politics and Reality*. Statistics Canada and US Census Bureau. Washington, D.C.: US Government Printing Office.
- Perlmann, J., and P. Nevada. 2015. “Ethno-Racial Origin in US Federal Statistics: 1980–2020.” Working Paper No. 857. Annandale-on-Hudson, N.Y.: Levy Economics Institute of Bard College. December.
- Skrentny, J. D. 2004. *The Minority Rights Revolution*. Cambridge, Mass.: Harvard University Press.
- US Census Bureau. 1986. *1980 Census of Population and Housing: Procedural History*. Washington, D.C.: US Government Printing Office.
- . 1996. *1990 Census of Population and Housing: Procedural History*. Washington, D.C.: US Government Printing Office.
- . 2003. “Oral History—Conrad Taeuber.” Interview of Conrad Taeuber conducted April 12, 1989, by Robert Voight. Available at www.census.gov/prod/2003pubs/oh-Taeuber.pdf.