The Leadership Conference

November 7, 2024

Shana Banks
Advisory Committee Branch Chief
Office of Program, Performance and Stakeholder Integration (PPSI)
U.S. Census Bureau
4600 Silver Hill Road
Washington, DC 20233

Submitted via email to shana.j.banks@census.gov

Re: Recommendations to the National Advisory Committee on Racial, Ethnic, and Other Populations on Coding of Race and Ethnicity Data under 2024 SPD 15 Standards

Dear Ms. Banks,

On behalf of The Leadership Conference on Civil and Human Rights, we applaud the Census Bureau for taking time at this point in the decade to discuss the coding and tabulation of American Community Survey (ACS) data beginning in 2027 and 2030 Census data during the November meeting of the National Advisory Committee on Racial, Ethnic, and Other Populations, given the recent revisions to Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity (SPD15). We believe that this early presentation will allow adequate time for the incorporation of stakeholder feedback, stringent testing, and the publication of findings before 2030 Census data collection begins.

The Leadership Conference is a coalition charged by its diverse membership of more than 240 national organizations to promote and protect the rights of all persons in the United States. The Leadership Conference Education Fund builds public will for state and federal policies that promote and protect the civil and human rights of all persons in the United States. Our coalition views an accurate and fair census, and the collection of useful, objective data about our nation's people, housing, economy, and communities generally, to be among the most important civil rights issues of our day.

Below, we offer several recommendations on coding, tabulation, and reporting of race and ethnicity data to support the successful implementation of the 2024 SPD 15 standards. These recommendations are focused in three main areas:

1. Engagement with stakeholders

2. Ensuring coding that matches respondents' intent
3. Transparent testing of coding and tabulation schemes

Engagement with stakeholders

As noted above, we are heartened that the Census Bureau is talking so early in the decade about tabulation, coding, and editing of data collected through the new combined question for race and ethnicity. The NAC presentation is a welcome first step in engaging key stakeholders in this consequential process, just as the Bureau consulted with the stakeholder community in equally important decisions about the minimum categories and questionnaire design.

We recommend that the Census Bureau engage not only with the academic and advocacy communities, but also with respondents – through qualitative research – to ensure that their intent is being captured in the Bureau's coding scheme, as we discuss below. This engagement is incredibly important to each of the following steps.

Ensuring coding that matches respondents' intent

The ability for everyone to write in more detailed information about each of the minimum categories for race/ethnicity reporting is a step forward for respondent self-identification. However, the new format can only live up to its full potential if the Census Bureau accurately records respondents' intent with respect to the use of checkboxes and write-in areas. Correctly discerning intent becomes even more important in light of potential confusion or uncertainty over the meaning of race, ethnicity, nationality, and national origin. In our highly interconnected world with complex histories of colonialism and migration, these concepts can mean vastly different things to different people. For the same reasons, old assumptions about the race of people indicating a country-specific national origin should no longer take precedence in coding and tabulation decisions. It is incumbent on the Census Bureau to ensure that question format, wording, and terminology steer respondents in the right direction with respect to selecting/providing responses to the race and ethnicity question that capture the categories with which they identify.

Examples from the 2020 Census can help to highlight this idea. According to the Census Bureau, the population reporting two or more races increased substantially between the 2010 and 2020 Censuses. Part of this increase was due, no doubt, to the growing diversity of the population of the United States. However, part of the growth of the "Two or more races" category was also likely due to a change in the question format that allowed respondents who selected the White and/or Black race categories to write in one or more detailed origins.

Overall, the new opportunity for respondents of all races to write in a subgroup in Census 2020 was a significant step towards ensuring more inclusive data collection and better self-identification. However, the Bureau's editing and coding decisions may have undermined the intent of some respondents to self-identify fully by race and subgroup or national origin.

Take, for example, a person who selected the Black or African American checkbox in the race question, and then wrote in "British" beneath the Black category. The coding rules for the 2020 Census assigned this person to both the Black and White race categories, thus adding them to the "Two or more races" category. Was this the respondent's intent? It might have been; but perhaps the respondent was born and raised in England and later moved to the United States, thereby identifying as British, but only as Black, not White, racially.

A second example highlights troubling coding and tabulation outcomes for responses to separate race and Hispanic Origin questions. A combined race and ethnicity question might not prevent such unintentional outcomes without careful development of coding and editing schemes for the 2027 ACS and 2030 Census, made in consultation with diverse external stakeholders and experts, such as members of the agency's advisory committees.

In this example, different outcomes would have emerged in 2020 depending on how a respondent filled out the two questions. An individual who views themselves as Hispanic and Black might have identified as Hispanic or Latino in the Hispanic Origin question and written in, for instance, "Dominican" as their specific Hispanic origin. Then, in the race question, they selected Black. The outcomes diverge depending on whether the respondent *also* wrote in "Dominican" under the Black race category: if the respondent did not write in a subgroup under that minimum category, the Census Bureau would have coded the respondent simply as Hispanic and Black. However, if this individual selected the Black race category and then wrote in "Dominican" underneath, the Census Bureau would place them, via coding, in the Hispanic, Black, *and Some Other Race* categories. In other words, if they provided fewer details, they were coded as Hispanic and Black, which is how they self-identified. Paradoxically, seeking to provide more detail by using the write-in space under the race category would have caused the Census Bureau to categorize this individual as Some Other Race, a category they did not select. While a combined race and ethnicity question might mitigate this unsupported outcome, we believe more logical, transparent coding, editing, and tabulation protocols going forward are the only way to capture respondent intent most accurately.

Transparent testing of coding and tabulation schemes

Both of the examples above show the importance of making respondent intent paramount in developing editing and coding guidelines. The Census Bureau cannot accomplish this goal without transparent testing of the coding and tabulation schemes. It is important to test the way data are coded and tabulated to understand the implications of any proposed changes in the protocols from one census to the next. The 2020 Census would have been a useful real-time opportunity to evaluate the potential effects of question format changes between the 2010 and 2020 enumerations. We hope the Bureau will create transparent testing opportunities to examine how changes in the collection of race and ethnicity data, as well as related coding and editing schemes for the 2030 Census, capture true demographic change rather than changes arising from new data collection and tabulation methods.

Additionally, the Census Bureau should test ways to determine the intent of respondents who use the write-in boxes, and how that intent compares with the editing and coding procedures that the Census

Bureau will use for the 2027 ACS (and beyond) and the 2030 Census. Publishing results from cognitive interviews to measure the reliability of the coding process, for example, would help all stakeholders better understand how both question design improvements and coding decisions bear on the Bureau's ability to capture respondent intent correctly.

Thank you for your consideration of our views on these important issues related to the collection and publication of race and ethnicity data. If you have any questions about these comments, please contact Meeta Anand, senior director of the census and data equity program at The Leadership Conference on Civil and Human Rights, at anand@civilrights.org.

Sincerely,

The Leadership Conference on Civil and Human Rights
The Leadership Conference Education Fund