Measuring the Racial and Ethnic Composition and Diversity of the U.S. Population: Historical Challenges and Contemporary Opportunities

By

Eric Jensen,[1] Nicholas Jones,[2] and Bev Pratt[3]

Population Division, U.S. Census Bureau

This presentation is released to inform parties of ongoing research and to encourage discussion in progress. Any opinions and conclusions expressed herein are those of the author(s) and do not reflect the views of the U.S. Census Bureau.

Abstract

[1] Eric Jensen is the Senior Advisor for Population Estimates and Coverage Measurement
[2] Nicholas Jones is the Director of Race/Ethnicity Research and Outreach
[3] Bev Pratt is the Senior Analyst for Race/Ethnicity Research and Outreach

Field Code Changed

Every decennial census since the very first in 1790 has included a question on race; however, the race question in every single decade has always been different in some shape or form. For example, a category for "Mexican" was included in the 1930 Census, but there was not a formal question about Hispanic origin (i.e., "ethnicity") on the full decennial census until 1980. Since 1980, the Hispanic origin question has also undergone changes each decade. The frequent modification to these questions reflects the difficulty inherent with measuring the fluid and complex concepts of race or ethnicity. In this proposed paper, we will review the historical challenges and contemporary opportunities with accurately measuring the racial and ethnic composition and diversity of the U.S. population. Specifically, we will provide an overview of the evolution of the race and ethnicity questions from the 1790 Census to the 2020 Census, review the U.S. Census Bureau's efforts to measure and analyze racial and ethnic diversity for the 2020 Census, and discuss potential plans for improving the measurement of these characteristics in data collections moving forward.

Extended Abstract

Every decennial census since the very first in 1790 has included a question on race; however, the design of the question has always been different, in some shape or form (U.S. Census Bureau 2021; Pratt et al. 2015). The history of measuring race in the decennial census reflects over two centuries of social, political, and economic change in the United States (Humes and Hogan 2009). The history of that one question—asked in various ways—follows the trajectory of our nation, encapsulating the era of slavery and emancipation, the incorporation of American Indians into the general U.S. population, the annexation of parts of Mexico, immigration from Asia and Latin America, the admittance of Alaska and Hawaii as the 49th and 50th states, the Civil Rights Movement, the growth of the foreign-born population, and the increase of the Multiracial/ethnic population.

A question on Hispanic origin first appeared on the full decennial census in 1980, after the first federal standards for race and ethnicity data were established in 1977 by the U.S. Office of Management and Budget's (OMB) Statistical Policy Directive No. 15. Statistical Policy Directive No. 15 was established mainly due to new responsibilities to enforce civil rights laws. Prior to that, a category for "Mexican" was included in the 1930 Census, and the 1970 Census included a question about Hispanic origin on the long-form sample questionnaire. A Hispanic origin question has remained on the census ever since 1980. There again, the question, has undergone changes with each decade. Note that the 1997 Statistical Policy Directive was revised in 1997 resulting in the *Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity* (SPD 15). All Federal information collections that collect information on race and ethnicity must adhere to SPD 15.

The frequent modification to the race and ethnicity questions reflects the difficulty inherent with measuring race and ethnicity in the United States. In the proposed paper, we will review the historical challenges and contemporary opportunities with accurately measuring the racial and ethnic composition and diversity of the U.S. population using the decennial census. Specifically,

Field Code Changed

we will provide an overview of the evolution of the race and ethnicity questions from the 1790 Census to the 2020 Census, review the Census Bureau's efforts to measure racial and ethnic diversity, and discuss potential plans for improving the measurement of these characteristics in data collections moving forward.

This paper will address two of the issues of interest listed in the Call for Papers. These include (1) "Current and historical collection of race and ethnicity data" and (2) "The future of terminology and communication of measures, such as retiring the constructs of *minority* and *majority* and introducing more informative measures and concepts, such as the *diversity index*." The Census Bureau's Race/Ethnicity Research Team has done extensive work on these topics, and we anticipate our paper will contribute to a robust discussion to help make the 2024 National Bureau of Economic Research a successful conference.

History of Measuring Race and Ethnicity

The evolution of measuring race and Hispanic origin in the decennial census reflects changes in U.S. population and the methods used to collect the census data over more than two centuries. The original 1790 Census included three response categories: "Free White Females and Males," "All Other Free Persons," and "Slaves." Fast forward to the 2020 Census and there were eight categories: White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, and Some Other Race. Because the 2020 Census included the option for respondents to select more than one race category, there were 63 possible combinations of race responses.

For much of the history of the U.S. decennial census, the primary method for taking the census was to have U.S. Marshals, and later Census Bureau enumerators, collect the census information (Anderson 2015). The traditional census approach, where the census is administered to respondents, usually relied on the personal observations of the enumerator when measuring race (Gauthier 2002). The 1960 Census was the first to use the mail-out approach where self-response was the primary method for collecting the Census data. This effectively instituted self-identification as the conceptual framework for reporting information about racial and ethnic identity. Today, the Census Bureau values and prioritizes self-identification when collecting data on race and Hispanic origin. In fact, the design of the race and ethnicity questions for the 2020 Census provided an opportunity for individuals from all communities, for the very first time, to "write-in" more detailed information for every one of the race and ethnicity categories, which allowed us to code the maximum amount of information on race for each respondent and all communities in the United States (Marks and Rios 2021).

After SPD 15 was revised in 1997, beginning with Census 2000, respondents could select more than one category in the *race* question. This allowed the Census Bureau to produce statistics on the racial distribution of the U.S. population for both single race groups ("race alone" concept), multiple race groups ("in combination" or "Multiracial" concept), and the maximum number of people reporting one or more race groups for each category ("race alone or in combination" concept). This change to measuring race has been critical for allowing respondents to fully self-

Field Code Changed

identify their racial identity and not limit their response to one racial category. Subsequently, the 2010 Census and 2020 Census have both allowed respondents to select more than one race category and the Multiracial population has grown with each census (Jones et al. 2021).

In the proposed paper, we will provide more information about the history of measuring race and Hispanic origin in the decennial census. In addition, we will provide specific examples of changes to the race and ethnicity questions and categories over time.

Measuring Racial and Ethnic Diversity

Over the years, the Census Bureau has used different approaches to describe the racial and ethnic diversity of the U.S. population. Beginning in 2008, the Census Bureau started using the "Majority/Minority" framework to describe geographic places where the non-Hispanic White population fell below 50 percent of the total population. This framework was applied to decennial census results, population estimates, and population projections. In particular, the Population Projections program received considerable attention for releasing their forecasts of when the "majority-minority crossover" would occur for the total U.S. population (Colby and Ortman 2015). However, this way of conceptualizing the changing composition of race and ethnicity was met with criticism by some race scholars (Alba 2017).

As we prepared to analyze and release the race and ethnicity results for the 2020 Census, the Census Bureau undertook concerted efforts to develop plans to ensure we presented information on race and ethnicity in ways that accurately reflected the characteristics of the population, used statistically appropriate methods, and showed respect to census and survey respondents who trust us with their information. In 2019, the Population Division of the Census Bureau formed the *Measuring Racial and Ethnic Diversity Working Group* to review contemporary research on race and ethnicity, consult with race and ethnicity scholars about the presentation of demographic statistics on race and ethnicity, and develop recommendations about how to present information on racial and ethnic diversity in the 2020 Census data products and beyond. The Working Group was made up of Census Bureau subject-matter experts in the areas of race and ethnicity, applied demography, statistical measurement, and data visualization.

One of the key issues that the Working Group discussed was the Census Bureau's past and future uses of the Majority/Minority framework. The Working Group came to a consensus that the agency would not use the Majority/Minority framework moving forward. The Working Group felt that it needed to be transparent about how the Census Bureau used the Majority/Minority framework in the past and why it discontinued its use.

In August 2021, based on the Working Group's analyses and data visualizations, the Census Bureau released several measures to describe the racial and ethnic diversity of the U.S. population in the 2020 Census. These included the diversity index, prevalence rankings and the diffusion score, and prevalence maps (Jensen et al. 2021).

Field Code Changed

In this paper, we will document how the Census Bureau came to adopt the Majority/Minority framework as a lens for measuring racial and ethnic diversity and then why we stopped using this approach in favor of other statistics to communicate racial and ethnic diversity. We will discuss how the measures mentioned above highlight more nuance and complexity to inform us about racial and ethnic diversity than the more simplified Majority/Minority framework.

The Future of Measuring Race and Ethnicity

The Census Bureau has made revisions and improvements to the race and ethnicity questions for the decennial census and demographic surveys since 1980 based on data-driven evidence from empirical qualitative and quantitative research, as well as consultation and engagement with stakeholders, advisors, and the public, and adhering to the U.S. Office of Management and Budget's (OMB) statistical standards for collecting and reporting race and ethnicity (OMB 2017). Future improvements to the measurement of race and ethnicity, and the design of a question or questions to do so, will be guided by these three factors – empirical research, public engagement, and Federal standards.

In 2015, the Census Bureau conducted the National Content Test (NCT) in preparation for the 2020 Census. The NCT included a combined race and Hispanic origin question to test if this format created less burden and provided better data than the two separate questions approach. The results showed that using a combined question format—with multiple detailed checkboxes and write-in fields and a dedicated for Middle Eastern or North African (MENA) respondents– reduced respondent burden and improved the accuracy of the results, especially for the Hispanic or Latino population and the multiracial/ethnic population (Matthews et al. 2017).

Another key finding from the NCT was that using the combined question format greatly reduced the number of people identifying as Some Other Race (SOR). SOR was intended to be a relatively small category but was the third largest race group in both the 2000 Census and 2010 Census (Humes et al. 2011) and has since grown to be the second largest race alone or in combination category in the 2020 Census, at 50 million people (Jones et al 2021). SOR is often selected by Hispanic or Latino respondents who do not see their racial identity reflected in the SPD 15 race categories. While not all Hispanic or Latino respondents identify as SOR, nearly all SOR responses are Hispanic or Latino, and reducing the size of the SOR race group was an improvement to the data collected in the 2015 NCT.

Ultimately, the greatest factor in changing the design of the race and ethnicity question(s) is OMB's 1997 SPD 15. The 1977 SPD 15 was revised in 1997 in part to respond to concerns that it did not adequately allow for self-identification, especially for people of Multiracial heritage, or reflect the increasing diversity of the U.S. population. In 2022, OMB reopened the review of SPD 15 by establishing an Interagency Technical Working Group (Working Group) to review the 1997 SPD 15and to develop recommendations for improving data on race and ethnicity (Orvis 2023). Per OMB's (2023) January 27, 2023 Federal Register Notice, the Working Group's initial proposals for revising SPD 15 include:

Field Code Changed

- Collecting race and ethnicity information using one combined question
- Adding “Middle Eastern or North African” (MENA) as a new response category
- Requiring the collection of detailed race and ethnicity categories by default

The Working Group is also asked for public comment on how best to revise SPD 15’s:

- Terminology, definitions, and question wording
- Guidance on implementation and adoption of revised standards
- Guidance on how to collect and report administrative and proxy data

The final decision by the OMB on whether and how to revise SPD 15 will be the biggest factor as to whether the Census Bureau implements a combined race and ethnicity question design with a distinct MENA reporting category for the 2030 Census. OMB has communicated publicly that it expects to decide by the Summer 2024. The initial proposals of the Working Group are grounded in empirical research such as, but not limited to, the 2015 NCT that found the combined question reduced respondent burden and improved data quality and accuracy for the decennial census and the American Community Survey.

In the proposed paper, we will provide more details on the work and findings of the Working Group’s review and proposals, and we will expand on the process for changing how race and ethnicity could be measured in the 2030 Census and how new data tabulations would inform our understand of the complexity and rich diversity of our Nation’s racial and ethnic composition.

Field Code Changed

References:

Alba, Richard. "What majority-minority society? A critical analysis of the Census Bureau's projections of America's demographic future." *Socius* 4 (2018).

Anderson, Margo J. *The American census: A social history*. Yale University Press, (2015).

Colby, Sandra L., and Jennifer M. Ortman. "Projections of the Size and Composition of the US Population: 2014 to 2060. Population Estimates and Projections. Current Population Reports. P25-1143." *US Census Bureau* (2015).

Gauthier, Jason G. *Measuring America: The decennial censuses from 1790 to 2000*. US Department of Commerce, Economics and Statistics Administration, US Census Bureau, (2002).

Humes, Karen, and Howard Hogan. "Measurement of race and ethnicity in a changing, multicultural America." *Race and Social Problems* 1:111-13 (2009).

Humes, Karen, Nicholas A. Jones, and Roberto R. Ramirez. "Overview of race and Hispanic origin: 2010." *US Census Bureau* (2011).

Jensen, Eric, Nicholas Jones, Kimberly Orozco, Lauren Medina, Marc Perry, Ben Bolender, and Karen Battle. "Measuring racial and ethnic diversity for the 2020 census." *US Census Bureau* (2021).

Jones, Nicholas, Rachel Marks, Roberto Ramirez, and Merarys Ríos-Vargas. "Improved race and ethnicity measures reveal US population is much more multiracial." *US Census Bureau*, August 12, 2021. (2021).

Marks, Rachel and Merarys Rios-Vargas. "Improvements to the 2020 Census Race and Hispanic Origin Question Designs, Data Processing, and Coding Procedures." *US Census Bureau*, August 3, 2021. (2021).

Matthews, Kelly, Jessica Phelan, Nicholas A. Jones, Sarah Konya, Rachel Marks, Beverly Pratt, Julia Coombs, and Michael Bentley. "2015 national content test: Race and ethnicity analysis report." *US Census Bureau* (2017).

Office of Management and Budget. "Initial Proposals for Updating OMB's Race and Ethnicity Statistical Standards." *88 Fed. Reg. 18, 5375-5384*. (2023).

Office of Management and Budget. "Standards for maintaining, collecting, and presenting federal data on race and ethnicity." *Statistical Policy Directive No 15* (1997).

Field Code Changed

Orvis, Karin. **“**Initial Proposals for Revising the Federal Race and Ethnicity Standards.” *Office of Management and Budget.* Accessed 3/21/2023. (2023) https://www.whitehouse.gov/omb/briefing-room/2023/01/26/initial-proposals-for-revising-the-federal-race-and-ethnicity-standards/

Pratt, Beverly, Lindsay Hixson, and Nicholas Jones. "Measuring Race and Ethnicity Across the Decades: 1790–2010." *Washington, DC: US Census Bureau. https://www. census. gov/data-tools/demo/race/MREAD_1790_2010. html* (2015).

U.S. Census Bureau. **“**U.S. Decennial Census Measurement of Race and Ethnicity Across the Decades: 1790–2020.” Data Visualization. Accessed 3/21/2023. (2021) https://www.census.gov/library/visualizations/interactive/decennial-census-measurement-of-race-and-ethnicity-across-the-decades-1790-2020.html

Field Code Changed