# Measuring the Racial and Ethnic Composition and Diversity of the United States Population: Historical Challenges and Contemporary Opportunities

Eric Jensen[1] and Nicholas Jones[2]

Population Division, U.S. Census Bureau

***Abstract***

Every decennial census since the very first in 1790 has included a question on race; however, the race question in every decade has always been different in some shape or form. For example, a race 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 to measuring the fluid and complex concepts of race or ethnicity. In this paper, we review the historical challenges and contemporary opportunities with accurately measuring the racial and ethnic composition and diversity of the U.S. population. Specifically, we 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 Federal data collections moving forward.

*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 authors and do not reflect the views of the U.S. Census Bureau.*

[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

## *Introduction*

Every decennial census in the United States, since the very first in 1790, has included a question on race. However, the design of the race 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 U.S. decennial census reflects over two centuries of social, political, and economic change in the United States (Humes and Hogan 2009). The history of the race 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/Multiethnic population.

A separate 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 race 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 these characteristics in the U.S. population. In this paper, we 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, we 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 addresses two of the issues of interest listed for the CRIW Conference. 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*."

We organize the paper in three sections. Section 1 discusses the history of measuring race and ethnicity in the United States through the lens of the U.S. decennial census, and the social context of the surrounding time periods. Section 2 describes work undertaken by the U.S. Census Bureau to measure and communicate analyses of racial and ethnic diversity with the 2020 Census results. Section 3 presents future opportunities for measuring race and ethnicity, with an update on work currently being undertaken in the Federal statistical system to review and revise SPD 15 to improve future data on race/ethnicity in the United States.

## Section 1. 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 six categories for race (White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, and Some Other Race) and two for ethnicity (Hispanic or Latino; Not Hispanic or Latino). Because the 2020 Census included the option for respondents to select more than one race category, there were 63 possible combinations of the race categories.

For much of U.S. decennial census history, 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 a "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, as people were asked to report their responses to the questions that were posed on the form.

Today, the Census Bureau values and prioritizes the concept of "self-identification" when collecting data on race and ethnicity. 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).

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). These changes to measuring race have been critical for allowing respondents to fully self-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).

### *History of Measuring Race and Ethnicity*

There are several excellent resources that discuss how the race and Hispanic origin questions on the decennial census have changed over time (Anderson 2015, Humes and Hogan 2009, Pratt et al. 2015). For example, Humes and Hogan (2009) outline five distinct historical periods in the measurement of race and ethnicity in the decennial census. Pratt et al. (2015) produced an infographic showing the specific race categories in each census from 1790 to 2010. Both use a

timeline approach that shows the evolution of the race and Hispanic origin questions over time. In this paper, we focus on each of the current racial and ethnic groups and highlight the key changes to the collection and tabulation for that group (Table 1).

The history of collecting information about the Black or African American population on the decennial census goes back to the first census in 1790, although these terms were not used. The census used the phrase "three-fifths of all other Persons," but it was clear to everyone that this refereed to the slave population (Humes and Hogan 2009). In 1820, the category "Free Colored People" was added to reflect the growth of the free-Black population. In 1850, blood quantum was introduced with the "Mulatto" as an attempt to capture the people with mixed heritage. This was expanded in 1890 to include the categories "Quadroon" and "Octoroon," which sought to quantify the percentage of Black blood a person had. These categories were dropped after one census because the data collected were not considered very accurate (Humes and Hogan 2009). However, "Mulatto" was added back as a category for the 1910 Census and the "one-drop" rule was added to enumerator instructions for the 1930 Census. The term "Negro" was first used on the Census in 1900 and remained on the census questionnaire until 2010 (Pratt et al. 2015). It was not until 2000 that "African American" was included as a category descriptor on the census.

The Asian category on the decennial census in many ways mirrors the historical trajectory of Asian immigration to the United States over the past two centuries. "Chinese" was added as a category to the 1870 Census following several decades of immigration from China (Humes and Hogan 2009). "Japanese" was added in 1890 after immigration from Japan increased when the U.S. government enacted laws restricting immigration from China (Table 1). For the 1920 Census, "Filipino," "Hindu," and "Korean" were added as race categories. "Hindu" and "Korean" were then dropped from the 1950 Census form, but "Korean" was added back on the form in 1970. "Asian Indian" and "Vietnamese" were added as race categories for the 1980 Census, the latter being added after the Vietnam War. Also in 1980, there were 6 detailed checkboxes added for the largest Asian Categories ("Japanese," "Chinese," "Filipino," "Korean," "Vietnamese," and "Asian Indian).

The 1980 Census was the first time that the different nationalities from Asia were combined to form the Asian category. The tabulations from the 1970 Census reported "White," "Negro," and "Other Race," which included "Indian," "Japanese," Chinese," "Filipino," and "All Other." The 1980 Census tabulated the Asian and Pacific Islander population, while still asking about the different Asian nationality groups on the form. It should be noted that the Census Bureau was following the then recently issued 1977 OMB Standards on Race and Ethnicity. Similarly, the Asian category was split out from the Pacific Islander group for the 2000 Census, again following the newly revised 1997 OMB Standards on Race and Ethnicity.

American Indians were not explicitly counted in the decennial census until 1860 when a category was added for "Indians" who are taxed (Table 1). The 1890 Census marked the first attempt to enumerate all American Indians (Humes and Hogan 2009). Blood quantum was dropped for the Black population in the 1900 Census but was still used on the American Indian questionnaire (Snipp 2000). With Alaska becoming a state in 1959, two new categories for "Eskimo" and "Aleut" were added to the 1960 Census form. Additionally, the term "American Indian" replaced

Table 1. Notable Milestones for the Decennial Census by Race and Ethnicity Category

| Race and Ethnicity | Census Year | Notable Milestones |
|---|---|---|
| Black or African American | 1790 | Slaves counted as “three-fifths of all other Persons” |
| | 1820 | "Free Colored Persons" added as a race category |
| | 1850 | Blood quantum introduced with the "Mulatto" category |
| | 1890 | ‘‘Quadroon’’ and ‘‘Octoroon’’ included as categories |
| | 1900 | The term Negro was used for the first time |
| | 1910 | “Mulatto” added back as a category |
| | 1930 | "One-drop" rule added to enumerator instructions |
| | 2000 | “African American” added as a category descriptor |
| | 2020 | “Negro” is removed from the race question |
| Asian | 1870 | "Chinese" added as a race category |
| | 1890 | "Japanese" added as a race category |
| | 1920 | “Filipino,” “Hindu,” and “Korean” added as race categories |
| | 1950 | “Hindu” and “Korean” dropped as race categories |
| | 1970 | “Korean” added back as a race category |
| | 1980 | “Asian Indian” and “Vietnamese” added as race categories and 6 detailed checkboxes added for the largest Asian categories |
| | 1990 | Data tabulated as Asian or Pacific Islander |
| | 2000 | Data tabulated as Asian after 1997 OMB Directive 15 |
| American Indian and Alaska Native | 1860 | Data collected on “Indians” who are taxed |
| | 1890 | First attempt to enumerate all American Indians |
| | 1900 | Blood quantum used to approximate assimilation |
| | 1960 | "American Indian" replaces "Indian" as a category; "Eskimo" and "Aleut" added after Alaska became a state |
| | 2000 | Data tabulated as American Indian or Alaska Native |
| Hispanic | 1930 | "Mexican" added as a race category |
| | 1970 | Hispanic self-identification question asked to a sample |
| | 1980 | Hispanic origin question asked to everyone and categories were added for specific Hispanic categories (Mexican, Puerto Rican, and Cuban) |
| Native Hawaiian or Other Pacific Islander | 1960 | "Hawaiian" and "Part-Hawaiian" added after Hawaii became a state |
| | 1990 | “Other Asian Pacific Islander" added as a category |
| | 2000 | New minimum category for Native Hawaiian or Other Pacific Islander; use of 3 detailed Pacific Islander checkboxes |
| Multiracial Population | 2000 | Census allows respondents to select more than one race |
| | 2020 | Expanded write-in and improved coding increases the size of the Multiracial population |
| Some Other Race | 1790-1840 | “All Other Free Persons” |
| | 1910-1990 | “Other”, “etc.?” and “Other Race” |
| | 2005 | “Some Other Race” is mandated by congress even though it does not follow the 1997 OMB standards |

Source: Adapted from Humes and Hogan (2009) and Pratt, Hixson, and Jones (2015).

"Indian" in 1960. Starting with the 2000 Census, the data for this population was tabulated as American Indian or Alaska Native.

In 1848, the Treaty of Guadalupe-Hidalgo was signed, and the United States annexed territory from Mexico that would become states in the current Southwest and West regions. Despite this change to the U.S. population, "Mexican" was not added to the census until 1930 when it became one of the race categories. A Hispanic self-identification question was added to the 1970 Census, but this was only one of the long-form items that were sent to a sample of households. Hispanic origin was included on the short-form of the 1980 Census. Also in 1980, categories were added for specific Hispanic categories: Mexican, Puerto Rican, and Cuban.

The categories "Hawaiian" and "Part-Hawaiian" were added to the 1960 Census after Hawaii became a state in 1959. "Other Asian Pacific Islander" was added as a category to the 1990 Census. At this time, the Native Hawaiian or Other Pacific Islander data were tabulated in the Asian Pacific Islander category. For the 2000 Census, Native Hawaiian and Other Pacific Islander became their own separate race category from Asian. Additionally, in 2000, three detailed Pacific Islander checkboxes for "Native Hawaiian," "Guamanian or Chamorro," and "Samoan" were added to the race question.

As noted above, the decennial census has a long and checkered history with collecting data on people with multiple racial identities. It seems that the underling purpose of asking about multiple racial identities for the Black population was exclusion, while the purpose for asking similar questions for American Indians was to measure assimilation and ultimately erase the population. The 2000 Census allowed respondents to select more than one race category and the Census Bureau tabulated the data as "alone or in combination" and "Two or more Races" categories. In contrast to past attempts to measure the Multiracial population, the goal of the new approach was to be inclusive of the growing number of people who identified and celebrated more than one race. The 2020 Census expanded the write-in categories and improved coding of the race responses, which allowed respondents to better report their complete heritages.

From 1790 to 1840, people who did not fit into the listed race groups were put into the "All Other Free Persons" category. For six decades there was not an "Other" category on the census. From 1910 to 1990, the category was labeled as "Other," "etc.?" and "Other Race." The 1997 OMB standards discouraged federal agencies from collecting an "Other" race category. However, Congress passed the 2005 Omnibus Appropriation Bill, which mandated that the Census Bureau must include "some other race" as a category when collecting census data. The reasoning behind requiring SORwas to provide a category for Hispanics who may not see themselves in the OMB categories (Humes and Hogan 2009). In the 2020 Census, SOR was the second largest alone or in combination race group in the United State—behind the White population—making up 15.1% of the population (Jones et al. 2021).

### *Comparative Approaches*

The changing approaches to measuring race and ethnicity in the census is not unique to the United States. Canada and Australia have followed similar historical paths to the United States

regarding how race and ethnicity have been categorized in the census data. However, in recent rounds of censuses the three countries take very different approaches (Stevens et al. 2015).

The United States, Canada, and Australia share similar historical, political, and social trajectories including their origin as a British colony, their status as immigrant receiving countries, and their large indigenous populations (Stevens et al. 2015). As former British colonies, the earliest collections of data on race and ethnicity focused mainly on establishing the size of the White population relative to all other groups.

This focus on the White population was evident in the immigration policies enacted by each country in the 19th and early 20th centuries. From 1882-1910, the United States enacted several laws that prohibited Chinese immigrants and their families from immigrating and becoming permanent residents (Calavita 2000). Canada adopted similar restrictive policies for Chinese immigrants (Holland 2007). Australian immigration policy long favored immigrants from the United Kingdom under a policy called "White Australia" that prevented immigration from other regions (Stevens and Fozdaar 2021). After World War II, these countries adopted new policies that allowed immigration from all over the world, which greatly increased the racial and ethnic diversity in each country.

A final similarity between these three countries is that they have large indigenous populations that were living in the territory at the time of contact and later became part of the country's population. For example, the United States did not include "Indian" as a category on the census until 1870, but since that time there has been a race category for the indigenous population on each subsequent census (Pratt et al. 2015). Canada's first census in 1871 included questions about "origins"; however, enumerators often faced resistance from the indigenous population when trying to conduct the census (Hamilton 2007; Kwan-Lafond and Winterstein 2020). Since Australia became an independent country in 1901, the aboriginal and Torres Straight Islander populations were included in the census, but they were not part of the official population totals until 1971 (Markham and Biddle 2018).

Despite these shared historical trajectories, today the United States, Canada, and Australia have significantly different approaches to measuring race and ethnicity in the census. The United States focuses on self-identified race and Hispanic origin for which respondents can choose between different options and write-in detailed information about their racial and ethnic identity. Since the 1996 census, Canada has used an approach to measuring race called the "visible minority" where respondents are asked "is this person…" and then gives a series of response categories including White, South Asian, Chinese, Black, Filipino, Arab, Latin American, Southeast Asian, West Asian, Korean, Japanese, and Other (Statistics Canada 2022a).

The responses are then categorized into visible minority status, which consists of "persons, other than Aboriginal peoples, who are non-Caucasian in race or non-White in color" (Statistics Canada 2022a). A separate question is used for indigenous status where people can identify as Inuit, First Nations, or Metis (Statistics Canada 2022b). Metis is a category that was originally used to identify the population in the Red River area of present-day Manitoba who had both indigenous and European ancestry, but the term has been used more broadly in recent decades by

people who identify as multiracial including indigenous origins (Kwan-Lafond and Winterstein 2020).

After the 1976 Census, Australia stopped including a race or racial origins question on their census; however, they still include variables such as religion, birthplace, parental place of birth, year of arrival, language spoken, indigenous status, and ancestry (Stevens et al. 2015). Public sentiment about the concept of "race" was the biggest factor in the decision to stop using a race question on the census, but the various indicators of social and cultural origins in Australia are still used as proxies for identifying different population groups.

### *Broader Limitations to the Census Approach to Measuring Race*

The Census Bureau's current focus on using self-identified race and ethnicity may have some limitations for measuring the impact of race on different socioeconomic outcomes. For example, research focusing on phenotype characteristics has found that wages, educational attainment, health outcomes, and other socioeconomic characteristics varies by skin tone (Dixon and Telles 2017; Goldsmith et al. 2006; Keith et al. 2017; Monk 2014). The empirical research shows a consistent pattern where lighter-skinned people have better outcomes than darker-skinned people.

Many of these studies focus on *within* group variation in phenotype characteristics. Monk (2021) argues that skin tone stratification within the Black or African American race group can sometimes be more pronounced than differences between the Black and White race categories. Studies have pointed to the variation in outcomes based on phenotype characteristics within a race group as an indicator of increased discrimination faced by people with darker skin tones (Keith et al. 2017). There can also be variation *between* race groups in phenotype characteristics that leads to increased discrimination, especially for the Hispanic or Latino population (Dixon and Telles 2017; Golash-Bonza and Darity 2008).

Some scholars have cautioned that measuring race and ethnicity using only self-identified categories may show different outcomes from methods that focus on phenotype characteristics (Lopez et al. 2018; Telles 2018). One novel method to overcome this problem combines the phenotypical approach and broad category approach by asking Hispanic or Latino respondents to not only report how they self-identify but also "how do you think others see your race" (Lopez et al. 2018). This research has found that the multidimensional approach to measuring race and ethnicity exposes more heterogeneity in health outcomes within the Hispanic or Latino population than just using a self-identified race and ethnicity question.

While it is unlikely that the Census Bureau will include questions about skin tone or perceived race on the decennial census or in large demographic surveys, these other approaches capture additional dimensions of stratification within race groups that cannot currently be captured with just self-identification.

### *Using Administrative Records*

With declining response rates to household surveys, organizations that collect data are increasingly using administrative records to populate race and ethnicity information for non-responding households (Ennis et al. 2018; Rastogi et al. 2017). The Census Bureau has conducted considerable research on this topic and even used administrative records to impute race and Hispanic origin information for non-responding households in the 2020 Census (Cantwell 2021). However, for this approach to provide valid data, the administrative records need to have accurate detail on race and ethnicity, which can be difficult because administrative data for these variables may be limited, outdated, or nonexistent.

Research looking at the correlation between self-reported race and race in administrative data has often found that the agreement varies across different race groups (Kressin et al. 2003; Ennis et al. 2018; Jarrin et al. 2020). Using data from the U.S. Department of Veteran Affairs, Kressin et al. (2003) found that there was the least agreement between administrative data and self-reported race for Native American, Asian, and Pacific Islander patients. Ennis et al. (2018) found that Hispanics were more likely to have non-matching race and ethnicity responses between administrative records and third-party data and their 2010 Census response. An analysis comparing race and ethnicity codes in Medicare administrative data to self-reported race found that the highest agreement was for the Black population and the lowest agreement was for the AIAN population (Jarrin et al. 2020).

Some administrative data sources do not include information about race and ethnicity, but researchers have developed methods for assigning these variables based largely on surname and geographic residence (Filice and Joynt 2017). Elliot et al. (2008) developed a Bayesian approach to estimate race and ethnicity using surname and geocoded residential addresses that they compared to self-reported responses. The model—the Bayesian Surname and Geocoding (BSG) method—was more efficient than a non-Bayesian method for predicting the individual's response and developing population-level statistics. Later research expanded this approach to develop the Bayesian Improved Surname Geocoding (BISG) method that uses additional information such as first and middle name to predict race in administrative data (Imai et al. 2022).

The U.S. Census Bureau has developed a hybrid approach where administrative data are linked to self-responses from a prior census or survey using probabilistic matching. The probabilistic matching process assigns a unique identifier called a Personal Identification Key (PIK) to individuals using the Social Security Administration Numerical Identification file (SSA Numident) as a reference file (Wagner and Layne 2014). The PIK is the foundation of the administrative records work being done by the U.S. Census Bureau.

The administrative records linked to past Census or ACS responses were used for the first time in the 2020 Census for item-nonresponse and the administrative records modeling portion of the non-response follow-up operation (Cantewell 2021). Table 1 shows that of the 331.4 million people in the 2020 Census, approximately 3.2% had their race and Hispanic origin response carried over from a past census or ACS response as part of the administrative records

Table 1. National Usage of Past Census and Administrative Records During Characteristic Imputation Processing: 2020 Census

| Characteristic | Percent of 331.4 million assigned from | |
|---|---|---|
| | Past Census or ACS response | Other Administrative Records |
| Race | 3.2% | 1.5% |
| Hispanic Origin | 3.2% | 0.9% |
| Age | 2.6% | 1.4% |
| Sex | 2.6% | 1.2% |

Source: U.S. Census Bureau, CBDRB-FY22-172

enumeration (Mule 2022). This was slightly higher than the age and sex characteristics where 2.6% of responses came from a past Census or ACS response. When a past census or ACS response was not available, we used information from other administrative records (Table 1). However, the information from other administrative records accounted for a relatively small proportion of the total demographic characteristics tabulated in the 2020 Census.

There are several concerns with this approach. First, if the PIK is not assigned correctly then the race response from a past census or survey response will not be valid. Even though research has shown that the false match rate for PIK assignment is relatively low (Layne et al. 2014), there is still a risk that the census or survey response is coming from another individual. Second, a person may change how they identify their race or Hispanic origin over time. Liebler et al. (2017) looked at matched Census 2000 and 2010 results and found that a non-trivial number of individuals changed their race and/or ethnicity response in the later census. Using administrative data linked to past census or survey responses will not capture potential changes in an individual's racial or ethnic identity.

Finally, the race and ethnicity data on the past census or survey will only be available in the format they were originally collected. This could be problematic if there are large changes to the way that data on race and ethnicity are collected on a future census or survey. For example, the past census and survey data linked to administrative records that were used for 2020 Census operations had race and ethnicity in the separate question format and may not align well with the data from a combined question approach. However, there could be resources developed to help map race and ethnicity data collected using different formats (see section on race bridging).

## Section 2. Measuring Racial and Ethnic Diversity

Over the years, the Census Bureau has used different approaches to describe changes in the racial and ethnic composition of the U.S. population. Beginning in 2008, we 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).

Changes to the racial and ethnic composition of the U.S. population are driven by subtle difference in demographic processes across groups. As birth rates decline for some groups and remain high for others, the racial and ethnic composition of the youngest cohorts will be different from older generations. Mortality can impact the racial and ethnic composition through differential mortality rates as well the age structure of the population. For example, the oldest-aged population, who are most at risk of dying, tend to be less racially and ethnically diverse than younger ages and therefore population decline from mortality will impact certain groups more than others. Immigration can also affect racial and ethnic composition. In recent decades, there has been substantial immigration from Latin America and Asia that has shifted the color line beyond the historical Black/White divide (Lee and Bean 2010).

Social scientists use different frameworks to describe the changing racial and ethnic composition of a population. These include the diversity index, segregation measures, and methods based on the percentage distribution of different groups (Johnson and Licther 2009; Iceland and Weinberg 2002; Massey, Rothwell, and Domina 2009; Meyer and McIntosh 1992). Each of these approaches highlight a different way to conceptualize and operationalize changes in the racial and ethnic composition of a population.

The diversity index is a probability-based measure of racial and ethnic diversity in a population (Meyer and McIntosh 1992). Conceptually, the diversity index measures how similar or different the racial characteristics of two people would be that are randomly selected from the population. The diversity index ranges from 0 to 1, with a 0-value indicating that the two people randomly selected from the population will have similar racial and ethnic characteristics, while a value of 1 indicates that two people randomly selected will have different characteristics. Additionally, there are other indexes for measuring racial and ethnic diversity, such as the Entropy Index.

Racial and ethnic segregation measures quantify the degree to which groups are concentrated or dispersed across different categories. Residential segregation focuses on where members of different racial and ethnic groups live within a specific geography (e.g., Metropolitan Statistical Area or county). Similar methods are used to measure occupational segregation where racial or ethnic groups are concentrated in a particular occupation or industry. Segregation is measured along different dimensions including evenness, isolation, clustering, centralization, and concentration (Massey and Denton 1988). Segregation measures highlight the amount of potential interaction between different race and ethnic groups within a specific geographic area.

As we released the results of the 2020 Census, it was important that we presented information on race and ethnicity in ways that accurately reflect the characteristics of the population, use statistically appropriate methods, and show respect to census and survey respondents who trust us with their information. In 2019, the Population Division of the Census Bureau formed the *Disseminating Diversity Working Group* to provide recommendations about how to present information on race and ethnic diversity in the 2020 Census data products and beyond. The Working Group is made up of 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. This approach has several conceptual and practical challenges that limit its ability to illustrate the complex racial and ethnic diversity of the U.S. population. For example, while some people classify individuals who identify with multiple population groups (such as Hispanic and White; White and Black or African American; and White and Asian) as part of the majority population, others classify them as part of the minority population. The dual identities of these groups highlight the social, political, and economic complexities of race and ethnicity in 21st century U.S. society.

The inclusion of certain groups as part of the "majority" or "minority" has also become more complex and contested in recent decades, especially as many people may not identify with certain population groups even if that is how they are classified and tabulated per SPD 15. The majority-minority approach is ambiguous, and it is further complicated by complex demographic and social realities.

To overcome these limitations, we focused on these alternative race and ethnicity diversity measures to illustrate the racial and ethnic composition of the 2020 Census results. 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 (as described above), prevalence ranking tables and graphs, and prevalence maps (Jensen et al. 2021).

### *Diversity Index*

Table 2 reports the states with the highest diversity index in the 2020 Census and the diversity index for these states in 2010. We converted the probabilities into percentages to make the results easier to interpret. In 2020, there was a 76 percent chance that two people chosen at random in Hawaii have a different race and ethnicity, which was a slight increase from 2010. Maryland was the state with the largest increase in the diversity index between 2010 and 2020 with a 6.6 percentage point increase.

Table 2. Ten States with the Highest Diversity Index: 2020 and 2010

| State | Diversity Index | | Percentage Point Difference |
|---|---|---|---|
| | 2010 | 2020 | |
| Hawaii | 75.1 | 76.0 | 0.9 |
| California | 67.7 | 69.7 | 2.0 |
| Nevada | 62.5 | 68.8 | 6.3 |
| Maryland | 60.7 | 67.3 | 6.6 |
| District of Columbia | 61.9 | 67.2 | 5.3 |
| Texas | 63.8 | 67.0 | 3.2 |
| New Jersey | 59.4 | 65.8 | 6.4 |
| New York | 60.2 | 65.8 | 5.5 |
| Georgia | 58.8 | 64.1 | 5.3 |
| Florida | 59.1 | 64.1 | 5.1 |

Note: Demographic changes, as well as improvements to the ways in which race and ethnicity data are collected and processed, reveal the United States population is more racially and ethnically diverse than measured in 2010.
States were selected using 2020 data and may not include the 10 states with the highest diversity index in 2010.

Percentage Point Difference based on unrounded values.
Information on confidentiality protection, nonsampling error, and definitions is available at https://www.census.gov/programs-surveys/decennial-census/technical-documentation/complete-technical-documents.html#redistricting

Source: U.S. Census Bureau, 2010 Census Redistricting Data (Public Law 94-171) Summary File. 2020 Census Redistricting Data (Public Law 94-171) Summary File.

### *Prevalence Rankings*

Prevalence rankings show the percent of the population that falls into the first-, second- or third-largest racial or ethnic groups. For 2020, the most prevalent racial or ethnic group for the United States was the White alone non-Hispanic population at 57.8%. This decreased from 63.7% in 2010. The Hispanic or Latino population was the second-largest racial or ethnic group, comprising 18.7% of the total population. The Black or African American alone non-Hispanic population was the third-largest group at 12.1%.

We also calculate the diffusion score, which measures the combined percentage of all racial and ethnic groups that are not in the first-, second- or third-largest racial and ethnic group. This calculation tells us how diverse and "diffused" the population is relative to the largest groups. The higher the score, the less concentrated the population is in the three largest race and ethnic groups. The remaining racial and ethnic groups combined to make up 11.4% of the total population in 2020, representing the diffusion score.

### *Prevalence Maps*

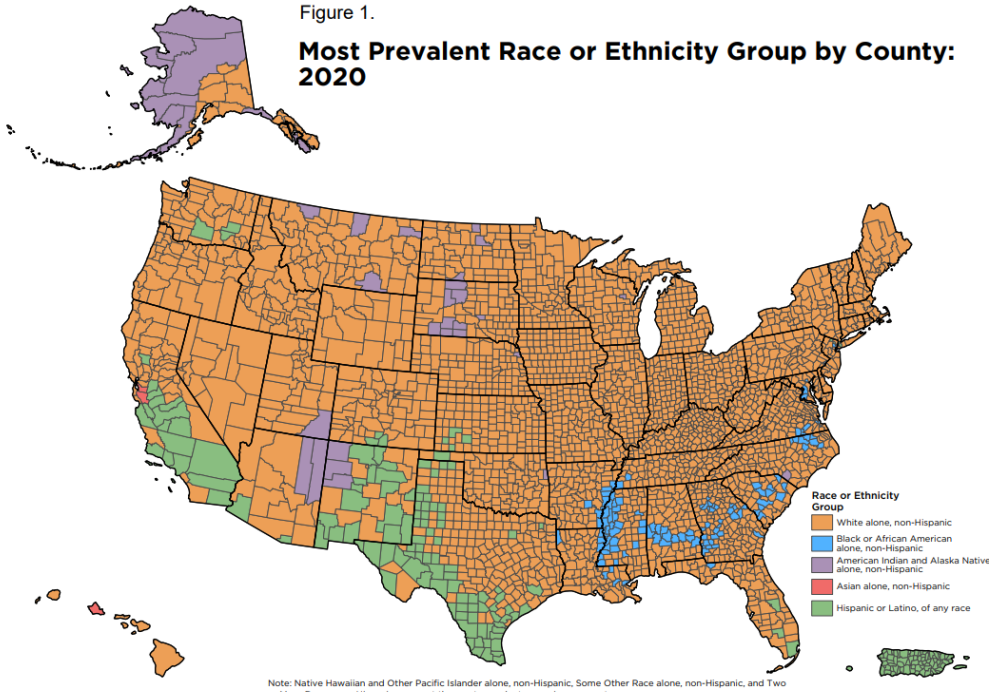
Prevalence maps show how the largest racial and ethnic groups are geographically distributed. The prevalence map for the most prevalent race or ethnicity group by county shows that for most counties, White alone non-Hispanic was the largest group (Figure 1). We do see some regional variations such as counties in the South where Black or African American alone non-Hispanic is the largest group. Also, there are quite a few counties in the Southwest where Hispanic or Latino is the largest race group. American Indian and Alaska Native alone non-Hispanic is the most prevalent group in some counties in the Great Plains, Four-Corners Area, and Alaska. Finally, Asian alone non-Hispanic was the largest in a few counties in California and Hawaii.

There is considerably more variation in the second-most prevalent race or ethnicity group by county (Figure 2). Large numbers of counties where the Hispanic or Latino population is the second-most prevalent group are found in every region, spanning the continental United States. The Multiracial non-Hispanic population was the second-most prevalent group in counties in the in the Midwest, along the northern border, an in Hawaii. The Black or African American population was the second-most prevalent race or ethnicity group in many of the counties in the South. Often, these patterns show an inverse relationship to the most prevalent group map.

Over the years, the U.S. Census Bureau has used different approaches to measure racial and ethnic diversity. For the 2020 Census results, we deliberately tried to move beyond limited approaches such as the Majority/Minority framework and focus on methods that highlight the complexity of racial and ethnic diversity in the United States. The measures that we have chosen—diversity index, prevalence rankings, prevalence maps, and diffusion scores—allow us to explore the full array of race and ethnic groups in a particular geographic area and not just the White/Other dynamic.

Figure 1.

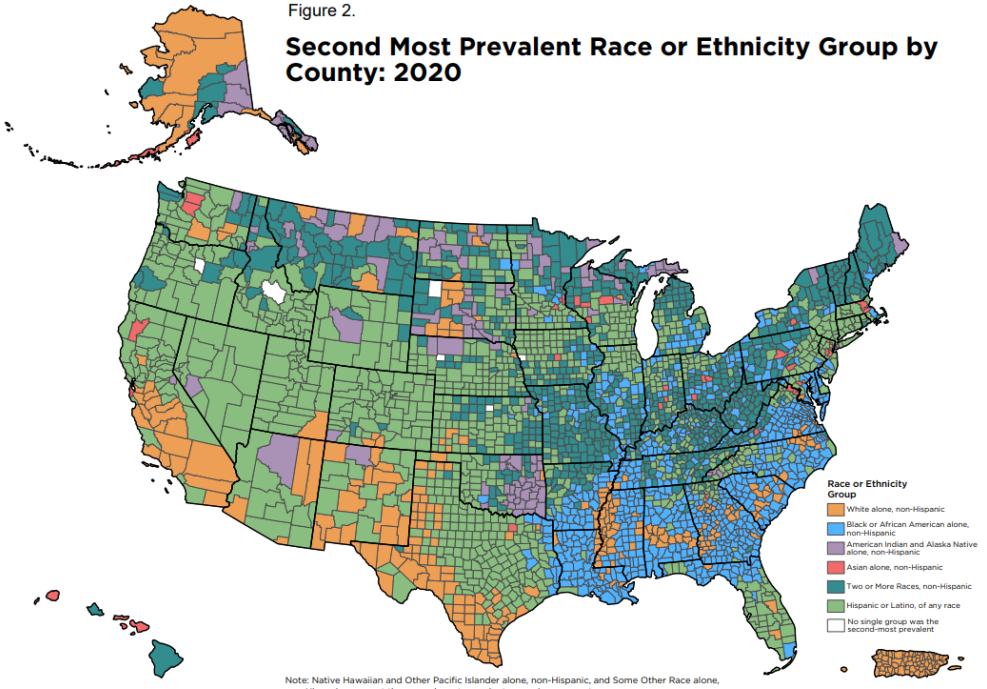
## Most Prevalent Race or Ethnicity Group by County: 2020




Note: Native Hawaiian and Other Pacific Islander alone, non-Hispanic, Some Other Race alone, non-Hispanic, and Two or More Races, non-Hispanic were not the most prevalent group in any county.
Source: Public U.S. Census Bureau, 2020 Census Redistricting Data (Public Law 94-171) Summary File.

Figure 2.

## Second Most Prevalent Race or Ethnicity Group by County: 2020




Note: Native Hawaiian and Other Pacific Islander alone, non-Hispanic, and Some Other Race alone, non-Hispanic were not the second most prevalent group in any county.
Source: U.S. Census Bureau, 2020 Census Redistricting Data (Public Law 94-171) Summary File.

## Section 3. Current Initiatives and Future Opportunities for Measuring Race and Ethnicity

### *Background: Previous Census Bureau Research on 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 evidence from empirical qualitative and quantitative research, as well as consultation and engagement with stakeholders, advisors, and the public, and adhering to OMB's SPD 15 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, write-in fields, and a dedicated option 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 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.

The U.S. Census Bureau has a long history of conducting research to improve questions and data on race and ethnicity. Since the 1970s, the Census Bureau has conducted content tests to research and improve the design and function of different questions, including questions on race and ethnicity. Over the past decade and a half, the Census Bureau conducted extensive research and outreach, including two groundbreaking national studies, on how to improve race and ethnicity question(s) so that these statistics better measure our nation's population, and to inform explorations of how Americans identify as our society continues to grow more diverse and more complex.

The Census Bureau's research last decade identified that a combined race and ethnicity question with multiple detailed checkboxes and a dedicated Middle Eastern or North African category is

the optimal design for improving race and ethnicity data, in comparison with designs which use two separate questions (Figure 3). This approach was strongly supported by myriad stakeholders and organizations.

However, the Census Bureau does not make a unilateral decision on the content of the Census. In fact, determining the content for a census is an extensive undertaking with a three-pronged approach involving empirical research, outreach, and engagement with stakeholders, and ultimately the review and approval from the U.S. Office of Management and Budget and the United States Congress.

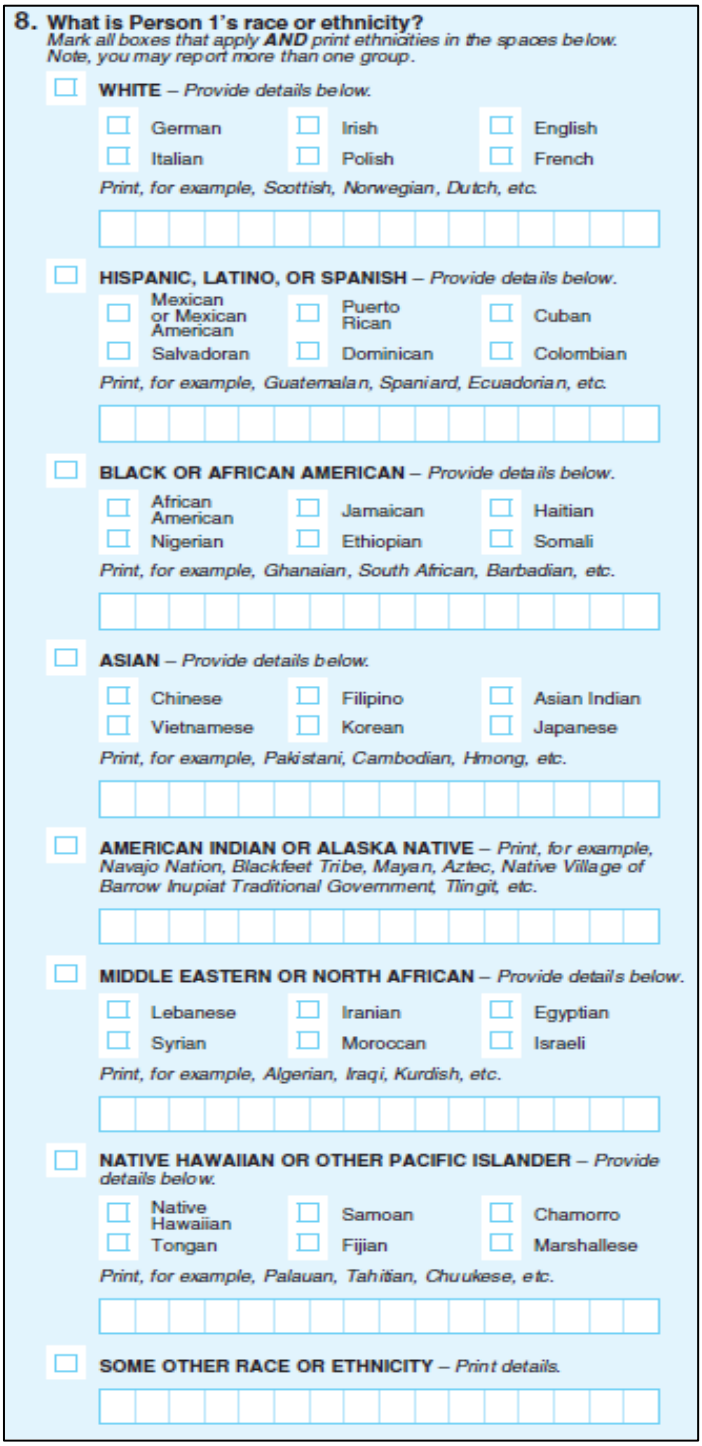
### *Race & Ethnicity in the 2020 Census*

In the 2020 Census, the Census Bureau collected race and ethnicity data in accordance with the 1997 OMB standards on race and ethnicity, and responses to these questions are based on self-identification. The Census Bureau, along with other Federal statistical agencies, must follow SPD 15, which require two separate questions when collecting data on race and ethnicity. The decennial census race question, for the first time in 2020, collected detailed responses for the White population and for the Black or African American population (Figure 4).

Our research and public feedback over the past decade illuminated strong interest from respondents to be able to self-identify their detailed racial/ethnic background, such as German, Lebanese, Mexican, Jamaican, Nigerian, Chinese, Navajo, Samoan, etc. SPD 15 encourages this collection of detailed responses, and to address this, new examples and write-in areas were added to the 2020 Census ethnicity question and race question to give respondents from all backgrounds the opportunity to self-identify their racial/ethnic identities in the 2020 Census. For the 2020 Census, we collected detailed responses for all major categories (Hispanic, White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, and Some Other Race). In turn, this provided the ability to produce detailed tabulations for myriad population groups in the United States, such as German, Lebanese, Mexican, Jamaican, Nigerian, Chinese, Navajo, Samoan, Brazilian, etc.

Building upon our research over the past decade, we improved the two separate questions design and updated our data processing and coding procedures for the 2020 Census. This work began in 2015 and was centered on findings from our National Content Test, with the new designs implemented in the 2018 Census Test. The improvements and changes enabled a more thorough and accurate depiction of how people self-identify, yielding a more accurate portrait of how people report their Hispanic origin and race within the context of a two-question format. These changes revealed that the U.S. population is much more multiracial and diverse than what we measured in the past.

## Figure 3. Optimal Design from Research Last Decade



**8. What is Person 1's race or ethnicity?**
*Mark all boxes that apply **AND** print ethnicities in the spaces below.*
*Note, you may report more than one group.*

☐ **WHITE** – *Provide details below.*

☐ German ☐ Irish ☐ English
☐ Italian ☐ Polish ☐ French

*Print, for example, Scottish, Norwegian, Dutch, etc.*

☐ **HISPANIC, LATINO, OR SPANISH** – *Provide details below.*

☐ Mexican or Mexican American ☐ Puerto Rican ☐ Cuban
☐ Salvadoran ☐ Dominican ☐ Colombian

*Print, for example, Guatemalan, Spaniard, Ecuadorian, etc.*

☐ **BLACK OR AFRICAN AMERICAN** – *Provide details below.*

☐ African American ☐ Jamaican ☐ Haitian
☐ Nigerian ☐ Ethiopian ☐ Somali

*Print, for example, Ghanaian, South African, Barbadian, etc.*

☐ **ASIAN** – *Provide details below.*

☐ Chinese ☐ Filipino ☐ Asian Indian
☐ Vietnamese ☐ Korean ☐ Japanese

*Print, for example, Pakistani, Cambodian, Hmong, etc.*

☐ **AMERICAN INDIAN OR ALASKA NATIVE** – *Print, for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Tlingit, etc.*

☐ **MIDDLE EASTERN OR NORTH AFRICAN** – *Provide details below.*

☐ Lebanese ☐ Iranian ☐ Egyptian
☐ Syrian ☐ Moroccan ☐ Israeli

*Print, for example, Algerian, Iraqi, Kurdish, etc.*

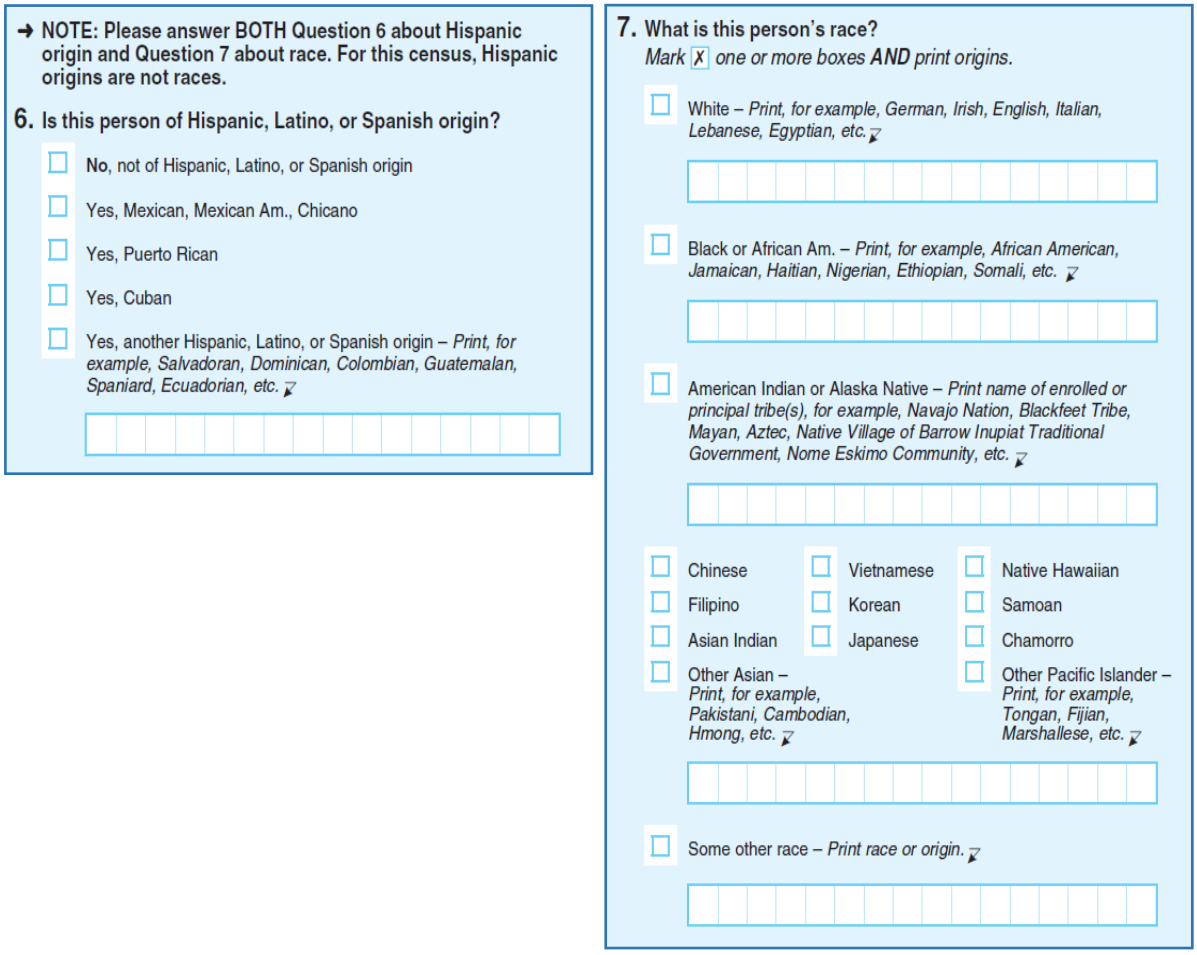
☐ **NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER** – *Provide details below.*

☐ Native Hawaiian ☐ Samoan ☐ Chamorro
☐ Tongan ☐ Fijian ☐ Marshallese

*Print, for example, Palauan, Tahitian, Chuukese, etc.*

☐ **SOME OTHER RACE OR ETHNICITY** – *Print details.*

**Figure 4. 2020 Census: Two Separate Questions Design**



➜ NOTE: Please answer BOTH Question 6 about Hispanic origin and Question 7 about race. For this census, Hispanic origins are not races.

6. Is this person of Hispanic, Latino, or Spanish origin?

- ☐ No, not of Hispanic, Latino, or Spanish origin
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban
- ☐ Yes, another Hispanic, Latino, or Spanish origin – *Print, for example, Salvadoran, Dominican, Colombian, Guatemalan, Spaniard, Ecuadorian, etc.*

7. What is this person's race?

*Mark ☒ one or more boxes **AND** print origins.*

- ☐ White – *Print, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.*
- ☐ Black or African Am. – *Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, 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.*
- ☐ Chinese
- ☐ Filipino
- ☐ Asian Indian
- ☐ Other Asian – *Print, for example, Pakistani, Cambodian, Hmong, etc.*
- ☐ Vietnamese
- ☐ Korean
- ☐ Japanese
- ☐ Native Hawaiian
- ☐ Samoan
- ☐ Chamorro
- ☐ Other Pacific Islander – *Print, for example, Tongan, Fijian, Marshallese, etc.*
- ☐ Some other race – *Print race or origin.*

However, results from the 2020 Census also showed the persisting problems with two separate questions on race and ethnicity. We are confident that differences in the overall racial distributions are largely due to improvements in the design of the two separate questions, as well as some demographic changes over the past 10 years. We are also confident, as shown in our research over the past decade, that using a single combined question for race and ethnicity in the decennial census would ultimately yield an even more accurate portrait of how the U.S. population self-identifies, especially for people who self-identify as multiracial or multiethnic.

### *OMB Review of Race Ethnicity Standards (SPD 15)*

OMB announced their formal review of SPD 15 in June 2022, with the goal of ensuring that the standards better reflect the diversity of the Nation. The Chief Statistician of the United States, Dr. Karin Orvis, and her staff organized the Federal Interagency Technical Working Group[3] (Working Group). The Working Group is undertaking a revision process similar to those used for the development and revision of other trusted statistical standards, which will help ensure the rigor, validity, objectivity, and impartiality of the resulting recommended revisions.

Consistent with OMB's established process, the Working Group is composed of Federal career staff who represent programs that collect or use race and ethnicity data from over 20 agencies across the federal government, including statistical agencies, almost every Chief Financial Officers Act agency, and the Equal Employment Opportunity Commission. The agencies on the Interagency Council on Statistical Policy, i.e., the 13 Principal Statistical Agencies[4], and the 24 agencies enumerated by the Chief Financial Officers Act[5], as well as the U.S. Equal Employment Opportunity Commission—selected for its reliance on race and ethnicity data—were invited to nominate representatives to the Working Group.

The Working Group has undertaken the important work of developing a set of recommendations for improving the quality and usefulness of Federal race and ethnicity data. The Working Group is evaluating relevant research, engaging in a meaningful way with the American public and all impacted agencies, and developing recommendations on topics including, but not limited to:

- Whether the minimum reporting categories should be changed and how to best address detailed race and ethnicity groups in SPD 15;
- Whether updates should be made to the question format, terminology, and wording of the questions, as well as the instructions for respondents and associated guidance; and
- Whether guidance for the collection and reporting of race and ethnicity data can be improved, including in instances when self-identification is not possible.

### *Census Bureau Leadership for Working Group*

The Census Bureau is working closely with OMB, serving as the Co-Chair of the Working Group, and providing subject matter expertise and guidance for each of the teams within the Working Group based on the Census Bureau's knowledge and previous research. Census Bureau experts are working with OMB and Federal agency colleagues to provide technical support and expertise for the formal review of SPD 15. The revised standards seek to improve the ability of the government to capture the diversity of the United States, to ensure that they are keeping pace with changes in the population and evolving needs and uses for data.

---

[3] Source: Information about OMB's Interagency Technical Working Group on Race and Ethnicity Standards <https://spd15revision.gov/content/spd15revision/en/about.html>.
[4] For reference, a list of the 13 Principal Statistical Agencies is available online <https://nces.ed.gov/FCSM/agencies.asp>.
[5] For reference, a list of the 24 Chief Financial Officers Act Agencies is available online <https://www.cfo.gov/about-the-council/>.

### *Development of Initial Proposals*

The Working Group reviewed existing research and evidence to develop initial proposals. For example, the Working Group reviewed what was done by the previous OMB Working Group last decade when they reviewed SPD 15, as they focused on many of the same issues that the current Working Group is tackling. Throughout, Census Bureau experts on race and ethnicity collaborated with fellow working group members and contributed to the discussions by presenting evidence from extensive research and engagement over the past decade, as well as data and experiences from the 2020 Census, and other sources. The initial proposals[6] were released in January 2023 for public comment, and are summarized below:

- **Initial Proposal 1:** Collect race and ethnicity information using one combined question.
- **Initial Proposal 2:** Add "Middle Eastern or North African" (MENA) as a new minimum category.
- **Initial Proposal 3:** Require the collection of detailed race and ethnicity categories by default.
- **Initial Proposal 4:** Update Terminology in SPD 15.
- **Initial Proposal 5:** Guidance is necessary to implement SPD 15 revisions on Federal information collections.

### *Race and Ethnicity Bridging*

As the standards for collecting data on race and ethnicity have changed over time, there is a need for resources to harmonize categories from one format to another. When the original 1977 standards were revised in 1997 this necessitated the need for bridging the two formats. The 1977 OMB race standards included four categories: American Indian or Alaska Native, Asian or Pacific Islander, Black, and White. In 1997, SPD 15 was revised with two major changes: 1) Pacific Islander became a separate category from Asian and 2) people could select more than one race category.

In Census 2000, for the first time, respondents were allowed to report multiple race categories. While this led to more accurate information and nuanced data about the complexity of people's racial identity and the composition of the U.S. population, having multiple race responses did not map on the race data collected by other state and federal agencies, many of whom were still using the 1977 OMB standards.

In 2003, the National Center for Health Statistics (NCHS) developed a series of adjustment factors that allowed data users to "bridge" from the 1977 OMB categories with only four single-

[6] For more information on the Initial Proposals for Updating OMB's Race and Ethnicity Statistical Standards, please visit <https://www.federalregister.gov/documents/2023/01/27/2023-01635/initial-proposals-for-updating-ombs-race-and-ethnicity-statistical-standards > (January 27, 2023).
**Note:** These proposals are preliminary and do not reflect the settled opinions of the Working Group, the position of OMB, or the positions of the agencies participating on the Working Group. The Working Group will continue to deliberate, assess evidence, and take into consideration comments received from the public before making final recommendations for OMB's consideration.

race categories to the 1997 OMB stands that included a possibility of 31 race categories (five single-race and 26 multiracial categories).

To develop the bridging factors, researchers at NCHS pooled data from the 1997 to 2000 National Health Interview Survey, which first asked respondents their race in the 1997 categories and then asked them to identify a "primary race" from the 1977 categories (Ingram et al. 2003). Next, they used logistic regression models to create the bridging factors from the 1997 to the 1977 categories. The models included covariates for demographic characteristics at the person level and contextual variables at the county level (Ingram et al. 2003, Leibler and Halpern-Manners 2008). The bridging had the biggest impact on the data for the AIAN and NHPI populations, which are the two groups with the largest proportion of multiracial individuals (Humes et al. 2011).

The NCHS bridging factors have been widely used by researchers to harmonize race data from the 1997 to the 1977 OMB standards. Additionally, researchers also need to convert data from the 1977 to the 1997 OMB standards, which the U.S. Census Bureau does using a process called "reverse bridging" (Sink and Colby 2014).

With the expectation that OMB will issue a revised SPD 15 by the Summer of 2024, members of the ITWG are doing research to develop bridging factors and reverse bridging factors that will allow data collected and tabulated using the 1997 OMB standards to be converted into a new standard, if one is adopted, that includes a combined race and ethnicity format. One option for this would be to use data from the 2015 National Content Test (NCT) conducted by the U.S. Census Bureau. The NCT included both a separate race and ethnicity format and a combined race and ethnicity format, similar to what has been proposed by the ITWG (Mathews et al. 2017). Once a revised SPD 15 is issued, it is expected that bridging will be one of the key topics that is undertaken to provide implementation guidance.

### *Census Bureau Race and Ethnicity Research and Testing Plans*

The Census Bureau's extensive research and engagement over the past decade and a half has shown that a combined race and ethnicity question with a Middle Eastern or North African (MENA) category and multiple detailed checkboxes is the optimal design for the decennial census and the ACS.

Census Bureau race and ethnicity subject matter experts are currently further exploring four key topics, building on previous research, and using 2020 Census data to fine tune knowledge on race and ethnicity question design. If OMB revises SPD 15, these research explorations will help inform and refine the optimal combined race and ethnicity question design with multiple detailed checkboxes and write-in areas for its use in the ACS and the 2030 Decennial Census. The four research questions are:

1. What are the item non-response rates for the two separate race and ethnicity questions?

2. What are the race reporting patterns of Hispanic or Latino respondents to the race question?

3. Within each minimum category, what are the frequencies of detailed population groups? What percentage do the six largest detailed groups comprise for each major category?

4. What do we know about reporting patterns of Middle Eastern or North African (MENA) responses in the 2020 Census race question?

The results from the additional research described above, in conjunction with OMB's decisions on the review of SPD 15, will be used to develop the plans for collecting and tabulating race and ethnicity in Census Bureau surveys including the ACS and the 2030 Census. Extensive experimentation and testing were conducted last decade and are not anticipated for the 2030 Census race and ethnicity questions. Should OMB revise SPD 15 to allow for a combined race and ethnicity question, there will be a need to be an operational field test to prepare for implementation in future Census Bureau data collections.

### *Next Steps*

The Census Bureau agrees with OMB that it is imperative for final revisions to be made and published by the summer of 2024. This is vital to ensure our programs and policies are effective across the Federal government. OMB is on track to reach the goal of completing these important revisions by the summer of 2024. The Working Group proposed that guidance be provided for implementing a revised SPD 15, and the Census Bureau expects to participate in these ongoing discussions.

Once a revised SPD 15 is issued by OMB, our Census Bureau race/ethnicity research leaders will review the updates and develop a plan to implement the new SPD 15 and its directives in our censuses and surveys, including the 2030 Census, the American Community Survey, and many other data collections. Until that time, we will not know OMB's decisions so we cannot speculate on what could or could not be possible with revised standards. As we develop initial plans, we expect to actively engage with stakeholders and the public to discuss the SPD 15 revisions and how they will inform and enhance our work and goals to improve race/ethnicity data for our nation.

### *Conclusion*

This paper contributes to the theme of the 2024 NBER conference in several ways. First, it provides historical context for how race and ethnicity have been measured in the decennial census. This history is not only about the groups that were included on the census form but also the groups that were not. Decisions on which groups to include or exclude were essentially about visibility, representation, and power. We also focus on the decision-making process that will determine how federal agencies collect, tabulate, and report information on race and ethnicity in the coming years, including the 2030 Census. The work of the ITWG is another milestone in the history of race and ethnicity in the United States outlined in this paper. Time will tell how well a revised SPD 15 will capture the complexities of race and ethnicity in contemporary American society.

## 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).

Calavita, Kitty. "The paradoxes of race, class, identity, and “passing”: Enforcing the Chinese Exclusion Acts, 1882–1910." *Law & Social Inquiry* 25, no. 1 (2000): 1-40.

Cantwell, Pat. "How we complete the census when households or group quarters don’t respond." *Random Samplings Census Blogs* (2021).

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

Dixon, Angela R., and Edward E. Telles. "Skin color and colorism: Global research, concepts, and measurement." *Annual Review of Sociology* 43 (2017): 405-424.

Elliott, Marc N., Allen Fremont, Peter A. Morrison, Philip Pantoja, and Nicole Lurie. "A new method for estimating race/ethnicity and associated disparities where administrative records lack self-reported race/ethnicity." *Health Services Research* 43, no. 5p1 (2008): 1722-1736.

Gauthier, Jason G. *Measuring America: The Decennial Censuses from 1790 to 2000*. U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau, (2002).

Golash-Boza, Tanya, and William Darity Jr. "Latino racial choices: The effects of skin colour and discrimination on Latinos’ and Latinas’ racial self-identifications." *Ethnic and Racial Studies* 31, no. 5 (2008): 899-934.

Goldsmith, Arthur H., Darrick Hamilton, and William Darity Jr. "Shades of discrimination: Skin tone and wages." *American Economic Review* 96, no. 2 (2006): 242-245.

Hamilton, Michelle A. "'Anyone not on the list might as well be dead': Aboriginal Peoples and the Censuses of Canada, 1851–1916." *Journal of the Canadian Historical Association* 18, no. 1 (2007): 57-79.

Holland, Kenneth M. "A history of Chinese immigration in the United States and Canada." *The American Review of Canadian Studies* 37, no. 2 (2007): 150.

Humes, Karen, Nicholas A. Jones, and Roberto R. Ramirez. "Overview of race and Hispanic origin: 2010." (2011): 1-23.

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.” *U.S. Census Bureau,* (2011).

Iceland, John, and Daniel H. Weinberg. *Racial and ethnic residential segregation in the United States 1980-2000*. Bureau of Census, (2002).

Imai, Kosuke, Santiago Olivella, and Evan TR Rosenman. "Addressing census data problems in race imputation via fully Bayesian Improved Surname Geocoding and name supplements." *Science Advances* 8, no. 49 (2022): eadc9824.

Ingram, Deborah D., Jennifer D. Parker, Nathaniel Schenker, James A. Weed, Brady E. Hamilton, Elizabeth Arias, and Jennifer H. Madans. "US census 2000 population with bridged race categories." (2003).

Jarrín, Olga F., Abner N. Nyandege, Irina B. Grafova, XinQi Dong, and Haiqun Lin. "Validity of race and ethnicity codes in Medicare administrative data compared to gold-standard self-reported race collected during routine home health care visits." *Medical Care* 58, no. 1 (2020): e1.

Jensen, Eric, Nicholas Jones, Kimberly Orozco, Lauren Medina, Marc Perry, Ben Bolender, and Karen Battle. "Measuring Racial and Ethnic Diversity for the 2020 Census." *U.S. Census Bureau*, (2021).

Jensen, Eric, Nicholas Jones, Megan Rabe, Beverly Pratt, Lauren Medina, Kimberly Orozco, and Lindsay Spell. "2020 US population more racially and ethnically diverse than measured in 2010." *Census Bureau* 12 (2021).

Johnson, Kenneth M., and Daniel T. Lichter. "Growing diversity among America's children and youth: Spatial and temporal dimensions." *Population and Development Review* 36.1 (2010): 151-176.

Jones, Nicholas, Rachel Marks, Roberto Ramirez, and Merarys Ríos-Vargas. "Improved Race and Ethnicity Measures Reveal U.S. Population is Much More Multiracial." *U.S. Census Bureau,* August 12, 2021, (2021).

Keith, Verna M., Ann W. Nguyen, Robert Joseph Taylor, Dawne M. Mouzon, and Linda M. Chatters. "Microaggressions, discrimination, and phenotype among African Americans: A latent class analysis of the impact of skin tone and BMI." *Sociological Inquiry* 87, no. 2 (2017): 233-255.

Kressin, Nancy R., Bei-Hung Chang, Ann Hendricks, and Lewis E. Kazis. "Agreement between administrative data and patients' self-reports of race/ethnicity." *American Journal of Public Health* 93, no. 10 (2003): 1734-1739.

Kwan-Lafond, Danielle, and Shannon Winterstein. "The Canadian census and mixed race: tracking mixed race through ancestry, visible minority status, and Métis population groups in Canada." *The Palgrave International Handbook of Mixed Racial and Ethnic Classification* (2020): 75-94.

Layne, Mary, Deborah Wagner, and Cynthia Rothhaas. "Estimating record linkage false match rate for the Person Identification Validation System." *Center for Administrative Records Research and Applications Working Paper* 2 (2014).

Lee, Jennifer, and Frank D. Bean. *The diversity paradox: Immigration and the color line in twenty-first century America*. Russell Sage Foundation, (2010).

Liebler, Carolyn A., and Andrew Halpern-Manners. "A practical approach to using Multiple-Race response data: A bridging method for public use microdata." *Demography* 45 (2008): 143-155.

Liebler, Carolyn A., Sonya R. Porter, Leticia E. Fernandez, James M. Noon, and Sharon R. Ennis. "America's churning races: Race and ethnicity response changes between census 2000 and the 2010 census." *Demography* 54, no. 1 (2017): 259-284.

López, Nancy, Edward Vargas, Melina Juarez, Lisa Cacari-Stone, and Sonia Bettez. "What's your "street race"? Leveraging multidimensional measures of race and intersectionality for examining physical and mental health status among Latinxs." *Sociology of Race and Ethnicity* 4, no. 1 (2018): 49-66.

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

Massey, Douglas S., and Nancy A. Denton. "The dimensions of residential segregation." *Social Forces* 67.2 (1988): 281-315.

Massey, Douglas S., Jonathan Rothwell, and Thurston Domina. "The changing bases of segregation in the United States." *The Annals of the American Academy of Political and Social Science* 626.1 (2009): 74-90.

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." *U.S. Census Bureau,* (2017).

Meyer, Philip, and Shawn McIntosh. "The USA Today index of ethnic diversity." *International Journal of Public Opinion Research* 4.1 (1992): 51-58.

Monk Jr, Ellis P. "The cost of color: Skin color, discrimination, and health among African-Americans." *American Journal of Sociology* 121, no. 2 (2015): 396-444.

Monk Jr, Ellis P. "The unceasing significance of colorism: Skin tone stratification in the United States." *Daedalus* 150, no. 2 (2021): 76-90.

Monk Jr, Ellis P. "Skin tone stratification among Black Americans, 2001–2003." *Social Forces* 92, no. 4 (2014): 1313-1337.

Mule, Thomas. "Quality of the Administrative Record Data Used in the 2020 Census." Presentation at the meetings of the Census Bureau's National Advisory Committee. May 6, (2022).

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).

Orvis, Karin. "Initial Proposals for Revising the Federal Race and Ethnicity Standards." *U.S. Office of Management and Budget*, (2023).

Pratt, Beverly, Lindsay Hixson, and Nicholas Jones. "Measuring Race and Ethnicity Across the Decades: 1790–2010." *Washington, DC: U.S. Census Bureau*, (2015).

Rastogi, Sonya, Leticia Fernandez, James M. Noon, Ellen Zapata, and Renuka Bhaskar. "Exploring administrative records use for race and Hispanic origin item non-response." (2017).

Snipp, C. M. American Indians and Alaska natives. In M. J. Anderson (Ed.), Encyclopedia of the US Census (pp. 28–31). Washington, DC: CQ Press. (2000).

Statistics Canada. “Indigenous Peoples Reference Guide.” Census of Population. (2022).

Statistics Canada. "Visible minority and population group reference guide." Census of Population (2022a).

Stevens, Gillian, Hiromi Ishizawa, and Douglas Grbic. "Measuring race and ethnicity in the censuses of Australia, Canada, and the United States: Parallels and paradoxes." *Canadian Studies in Population [ARCHIVES]* 42.1-2 (2015): 13-34.

Sink, Larry D., and Sandra L. Colby. “Race Bridging: An Evaluation of the Census Bureau’s Current Method for Converting NCHS Data into OMB 1997 Categories.” Paper presented at the 2014 Southern Demographic Association. Memphis, TN. (2014).

Telles, Edward. "Latinos, race, and the US census." *The ANNALS of the American Academy of Political and Social Science* 677, no. 1 (2018): 153-164.

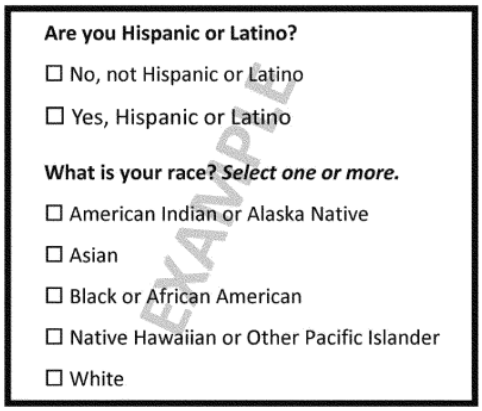
U.S. Census Bureau. “U.S. Decennial Census Measurement of Race and Ethnicity Across the Decades: 1790–2020.” Data Visualization, (2021).

Wagner, Deborah, and Mary Lane. *The person identification validation system (PVS): applying the Center for Administrative Records Research and Applications’(CARRA) record linkage software*. No. 2014-01. Center for Economic Studies, US Census Bureau, (2014).

## Appendix. Images from ITWG Initial Proposals for Revised Question Format

The Working Group proposes that SPD 15 move from the two separate questions format to a single combined question as the required design for self-reported race and ethnicity information collections. For more information on the initial proposals, please visit the website spd15revision.gov.

**1997 SPD 15's Two-Questions Format for Self-Response**




**ITWG Initial Proposed Example for Self-Response Data Collections: Combined Question with Minimum Categories**



**ITWG Initial Proposed Example for Self-Response Data Collections: Combined Question with Minimum and Detailed Categories**

**What is your race or ethnicity?**
*Select all that apply **AND** enter additional details in the spaces below.*
*Note, you may report more than one group.*

☐ **WHITE –** *Provide details below.*

☐ German ☐ Irish ☐ English
☐ Italian ☐ Polish ☐ French
*Enter, for example, Scottish, Norwegian, Dutch, etc.*

☐ **HISPANIC OR LATINO –** *Provide details below.*

☐ Mexican or Mexican American ☐ Puerto Rican ☐ Cuban
☐ Salvadoran ☐ Dominican ☐ Colombian
*Enter, for example, Guatemalan, Spaniard, Ecuadorian, etc.*

☐ **BLACK OR AFRICAN AMERICAN –** *Provide details below.*

☐ African American ☐ Jamaican ☐ Haitian
☐ Nigerian ☐ Ethiopian ☐ Somali
*Enter, for example, Ghanaian, South African, Barbadian, etc.*

☐ **ASIAN –** *Provide details below.*

☐ Chinese ☐ Filipino ☐ Asian Indian
☐ Vietnamese ☐ Korean ☐ Japanese
*Enter, for example, Pakistani, Cambodian, Hmong, etc.*

☐ **AMERICAN INDIAN OR ALASKA NATIVE –** *Enter, for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Tribal Government, Tlingit, etc.*

☐ **MIDDLE EASTERN OR NORTH AFRICAN –** *Provide details below.*

☐ Lebanese ☐ Iranian ☐ Egyptian
☐ Syrian ☐ Moroccan ☐ Israeli
*Enter, for example, Algerian, Iraqi, Kurdish, etc.*

☐ **NATIVE HAWAIIAN OR PACIFIC ISLANDER –** *Provide details below.*

☐ Native Hawaiian ☐ Samoan ☐ Chamorro
☐ Tongan ☐ Fijian ☐ Marshallese
*Enter, for example, Palauan, Tahitian, Chuukese, etc.*

EXAMPLE