Mathematica®
Progress Together

Response to the Request for Comments for

Initial Proposals for Updating Race and Ethnicity Statistical Standards

April 19, 2023

Submitted to:

Interagency Technical Working Group on Race and Ethnicity Standards
1650 17th St. NW
Washington, DC 20500
Attention: Bob Sivinski, Chair

Submitted by:

Mathematica
1100 First Street, NE, 12th Floor
Washington, DC 20002-4221
Phone: (202) 484-9220
Fax: (609) 228-4958

This page has been left blank for double-sided copying.

For more than 50 years, Mathematica has worked closely with federal agencies, grantees, foundations, community organizations, and other partners in the evidence community to collect, analyze, and interpret data on a wide range of topics, including health, education, labor, and social policies. As data experts, we recognize the distinct role the federal government holds in the evidence community and the importance of accurate and comprehensive data in promoting equity and justice for all people. We understand the complexities and nuances involved in collecting and analyzing data on race and ethnicity and believe strengthening the federal government's ability to make data-informed policy decisions is critical to advancing equity. We are grateful for the opportunity to contribute to this important dialogue and to support initiatives that seek to improve how we collect data on race and ethnicity in the United States.

We acknowledge that race and ethnicity are multidimensional constructs that vary across time and context. The complexities of race, ethnicity, and identity in the United States shift as migration patterns, social categories, and terminology evolve. In addition, personal connections to ancestry, ethnic groups, and national or regional heritage can change as people study their family histories through conversations with relatives, genealogical research, and commercially available DNA testing. Thus, incorporating response options that allow people to more fully describe their identities can enhance the overall validity of the data, as respondents might become skeptical of survey questions that do not provide them with an appropriate answer. Although perfection might be impossible, using evidence, data, and current trends to inform decisions and select survey items is critical for mitigating selection bias. It also increases visibility for historically marginalized communities that might be underrepresented in mainstream discourse, allowing recognition of their unique needs and contributions.

With this in mind, Mathematica fully supports the Federal Interagency Technical Working Group on Race and Ethnicity Standards (Working Group) in its efforts to revise and expand the federal race and ethnicity standards to better represent the changing identity, growing breadth, and increasing complexity of population groups in the United States. By accurately representing the diversity of this nation, we enhance our ability to make informed decisions about policies and programs that impact communities, ultimately advancing equity for all. For any questions about our response, please contact Mike Burns, Senior Director for Communications and Public Affairs, at MBurns@mathematica-mpr.com.

Question 1: Collect Race and Ethnicity Information Using Combined Question

Question 1.a - Please provide links or references to relevant studies that examine or test any impacts of collecting race and ethnicity information using separate questions compared to a combined question.

We have included references to relevant studies as embedded links throughout our responses to the specific questions.

Question 1.b - To what extent would a combined race and ethnicity question that allows for the selection of one or more categories impact people's ability to self-report all aspects of their identity?

Combining race and ethnicity in one question might better reflect the integrated view of racial and ethnic identity held by many Hispanic or Latino[1] people, resulting in fewer missing or invalid responses, or responses that indicate "some other race." A Pew Research Center report suggests many Hispanic adults do not distinguish between race and ethnicity as distinct concepts; in fact, two-thirds of Hispanic adults see their Hispanic background as part of their racial background. In a study by the U.S. Census Bureau, more than 40 percent of people who identified as Hispanic or Latino did not report belonging to any federally recognized race group as defined by OMB, when asked to identify their race and ethnicity separately. In response, the U.S. Census Bureau tested various "strategies to increase reporting within the major OMB race and ethnicity categories," including the 2010 Census Alternative Questionnaire Experiment and 2015 National Content Test. Both experiments found that a combined race and ethnicity question increased reporting within OMB categories, resulting in more respondents identifying as White or Hispanic, and fewer respondents identifying as "some other race" or providing missing or invalid responses.

Although combining race and ethnicity in one question would increase reporting in the major OMB race and ethnicity categories, it might not fully capture the multidimensional nature of Hispanic or Latino identity and enable people to self-report all aspects of their identities. According to the 2014 National Survey of Latinos, when asked directly whether they consider themselves to be "mixed race or multiracial," about one-third of Latino adults identified as mixed race, a rate more than five times higher than the share of Latinos who selected two or more races in responding to the standard race question. Similarly, a quarter of Latino adults identified as having Indigenous or Native American roots, compared with only 2 percent responding to the standard race question. Finally, a quarter identified as having Afro-Latino origins, compared with just 8 percent who indicated on the standard question that one of their races was Black. Combining race and ethnicity in one question using the standard OMB race categories would not address the limitations of the existing race categories to capture how Hispanic or Latino people identify themselves, and would continue to mask important subgroups, such as Afro-Latino and Indigenous Latino. However, results of the 2015 National Content Test suggest that combining race and ethnicity in one question would not exacerbate lost race information, as Hispanic or Latino respondents were no less likely to report their race as White or Black with combined race and ethnicity questions than with separate questions, and in fact were better able to find themselves among the race and ethnicity categories.

Question 1.c - If a combined race and ethnicity question is implemented, what suggestions do you have for addressing challenges for data collection, processing, analysis, and reporting of data?

As discussed in How Administrative Data Collection and Analysis Can Better Reflect Racial and Ethnic Identities, as people age and navigate life experiences and social contexts, their self-reported racial and ethnic identity can change, shaped by personal histories, the identities they disclose to

---

[1] In our response to questions, we use the term "Hispanic or Latino," as this reflects the current way Federal agencies have collected data on this diverse population. However, we acknowledge that these terms may not be fully inclusive, and some groups may prefer the use of gender-neutral alternatives, such as "Latinx." The Inclusive Language Guidelines provided by the American Psychological Association suggests consulting with study participants to determine the appropriate term when feasible. We recognize that language is dynamic and constantly evolving and language preferences may vary among different communities and individuals.

others, and what they believe others perceive them to be. People are also more likely to change their responses over time when they can select all that apply. This option offers a more valid measure of current racial and ethnic identity but a less reliable measure for assessing trends over time. To support validity and reliability of the data given this challenge, we recommend OMB take the following steps:

- *List clear instructions to help respondents provide accurate information, which can also help agencies and data users interpret the data.* Survey instructions should clarify the type of racial and ethnic identification the survey is asking about and, potentially, which types the survey is not asking about (for example, what you consider yourself to be, not what you believe others perceive you to be).

- *Ensure accurate and insightful data by carefully considering the design and flow of the questions used for collecting information on Hispanic or Latino racial and ethnic identity.* Question design should take into account research on the dimensions of <u>Hispanic or Latino racial and ethnic identity</u> to create a clear and comprehensive question flow for respondents. For instance, people who identify as Afro-Latino or Afro-Mexican might find it confusing to first identify as Hispanic Mexican, and then as Black.

There is a lack of consensus in the field about how to analyze race and ethnicity data when people identify with multiple groups, and a combined race and ethnicity question raises concern about whether Hispanic or Latino should (still) be reported as a mutually exclusive category from race. A common research practice is to group individuals who select more than one category into a "multiracial/ethnic" category, but this approach leads to difficulty with interpretation. Data from the <u>Pew Research Center</u> show that this group is highly heterogeneous even when focusing only on racial identification—for example, encompassing those who identify as both Black and American Indian and those who identify as both White and Asian. In addition, many people who select multiple races and ethnicities do not consider themselves "multiracial." A recent <u>Brookings research report</u> describes the implications of this issue, noting that "just 23% of all Native Americans were classified as single race and non-Latino or Hispanic on the 2020 Census" thus "excluding the more than three-quarters of the total Native American population" (Maxim et al. 2023) who selected more than one category. Another common research practice when people identify with multiple groups is to exclude their data from analysis, to avoid treating this group as a monolith and because there might be too few cases to statistically analyze meaningful intersectional groups (such as Afro-Latinos). To promote consistency and action-oriented insights despite these challenges, we recommend OMB take the following steps:

- *Develop clear guidance on how to process, analyze, and report information from a combined race and ethnicity question that encourages meaningful inclusion of data on people who identify with more than one category.* We recommend OMB allow communities to drive the guidance on how to combine, separate, and interpret categories and how to clean open-ended responses to define subcategories. We suspect these subcategories would include intersectional identities, such as Afro-Latino and Indigenous Latino rather than a monolithic Hispanic or Latino group. This supports the recommendation made by <u>Maxim et al. 2023</u> to "encourage federal government agencies to publish data on American Indian and Alaska Native populations along and in combination with other groups" expanding to include other intersectional identities.

- *Develop separate guidance for analyzing trends over time that includes data from years before surveys offered a combined race and ethnicity question.* In these cases, we recommend OMB encourage reporting of data in multiple ways—for example, treating Hispanic or Latino as a mutually exclusive category from race to describe trends over time for this group, while contextualizing those patterns with

differences within the Hispanic or Latino group using more detailed information from recent years, or supplementing responses with qualitative information.

Question 1.d - What other challenges should we be aware of that respondents or agencies might face in converting their surveys and forms to a one question format from the current two-question format?

Changing the format of data collection on race and ethnicity could create challenges for agencies that track and aim to maintain trends over time—for example, if some differences in distributions are due to changes in the question format or the detailed response categories, or if they must meet state or other regulatory requirements that specify a two-question format.

Agency staff will likely need to review the new question format and the reasons behind it, so they are prepared to talk with clients and respondents about the change, and to ensure all people use and report the data appropriately. This process might require educating staff about the history of race and ethnicity in the United States, its measurement, the benefits of the new format, and best practices for collecting this information in multimode studies that incorporate a telephone administration option.

Question 2: Add "Middle Eastern or North African" (MENA) as a new minimum category

Mathematica supports adding Middle Eastern or North African (MENA) as a new minimum category to questions of racial and ethnic identity in federal surveys, including Census Bureau products. This is a positive and important step toward addressing the longstanding concerns articulated by U.S. residents of MENA descent, community organizations, researchers, and government representatives. These groups have highlighted the harms of excluding this category and of requiring those who identify ethnically with MENA to select White when responding to surveys and questionnaires. Arab American organizations and researchers who specialize in MENA populations in the United States have called for the inclusion of a MENA category to represent their communities Abuelezam et al. 2018).

Excluding MENA residents from a separate category amounts to erasing their unique experiences and expanding the definition of White to include populations, experiences, and backgrounds that are not typical of those of primarily European ancestry. Importantly, having a separate MENA category will advance research and a better understanding of the social, educational, health, and economic factors and outcomes unique to MENA populations in the United States, which represent more than 30 countries across the region and myriad ethnic and cultural groups.

Question 2.a - Given the particular context of answering questionnaires in the U.S. (e.g., decennial census, Federal surveys, public benefit forms), is the term "Middle Eastern or North African (MENA)" likely to continue to be understood and accepted by those in this community? Further, would the term be consistently understood and acceptable among those with different experiences, i.e., those born in the U.S., those who immigrated but have lived for an extensive period of time in the U.S., and those who have more recently immigrated to the U.S.?

We believe the MENA community will understand and accept the proposed term, despite some disagreement about its use and the definition of regional boundaries. Some respondents might feel

the term is too broad and diminishes their personal identity in the name of inclusivity. However, similar to using Hispanic or Asian to capture the identities of people from wide geographic areas with significant differences in culture, migration patterns, language, and experiences, MENA is an imperfect but necessary category.

If possible, we recommend OMB clearly define the MENA category with geographic boundaries that resonate with MENA populations (even if there is some division about their exactness). Based on the most common definitions, the geographic boundaries encompassing this category would include Turkey but exclude Afghanistan or any country east of Iran. It might include Sudan and South Sudan (which are considered Arab countries). By defining the category with clear geographic boundaries, we believe that respondents will have a clear and accurate understanding of when to select MENA.

The Census' own empirical research shows that when no MENA category was present, about 86 percent of MENA respondents identified as White. However, when a MENA option was added to a list of racial and ethnic categories, only 20 percent of MENA respondents chose White, and the remaining 80 percent selected MENA ([Matthews et al. 2017](#)). The proportions of respondents identifying as MENA might be even higher today given that the inclusion and adoption of the term MENA has gained national attention as a category for these populations in federal surveys ([NPR, 2022](#)).

Question 2.b - Do these proposed nationality and ethnic group examples adequately represent the MENA category? If not, what characteristics or group examples would make the definition more representative?

The proposed groups within the MENA category might not be the most effective way to capture the full experience of the MENA population in the United States. While decisions on group examples should be data driven, there is currently no strong evidence or agreed upon approach in the literature. Notably absent is the option of Arab, a pan-national identity with substantial shared history, culture, and language. The [2021 American Community Survey](#), shows that Arab is by far the largest group of MENA-identifying people in the U.S., comprising more than 2 million residents. In fact, when MENA is examined by subcategories, 210,000 people identify as simply Arab rather than reporting a specific national ancestry.

Given the lack of an evidence-based approach to define or clarify MENA, we suggest that the Working Group consider the following options:

- Regional group examples (North Africa, Gulf, Levant/Sham, etc.) in addition to a few nationalities, or instead

- Focus on cultural, linguistic, and ethnic groups, particularly those with the largest populations in MENA, such as Arab, Persian, Kurdish, Turkish, non-Arab Israeli, Amezigh/Berber, and Assyrian/Chaldean/Syriac, and remove all consideration of country of origin. One limitation of this approach is the inability to disaggregate the very large Arab population to smaller groups represented in the previous example or by nationality.

Question 2.c - Would this proposed definition allow the generation of statistics necessary to track the experience and wellbeing of the MENA population?

The proposed definition would enable the collection of data necessary for monitoring and analyzing the unique experiences and well-being of the MENA population in the United States. Research conducted by the Arab American Institute suggests a substantial underrepresentation of Arab American populations in Census data. Importantly, the proposed group examples put forth will impact our ability to study relevant subgroups within the MENA category. The suggested group examples might serve as a more accurate representation of the MENA population compared to treating MENA as a single category. This approach is likely to yield more robust findings, particularly if the group examples are representative of MENA subgroups that have a significant presence in the United States.

Question 3: Require the collection of detailed race and ethnicity categories by default

We recommend not requiring the collection of detailed race and ethnicity categories, by default, on all federal collections. Instead, the decision to collect this more granular data should be grounded in the study's objectives, survey design, sample size, data collection methods, and analytic approach. Guidance that can help agencies effectively balance the need for the detailed categories with the concerns of survey quality and participant burden is needed. Such guidance would allow researchers and agencies to determine, in a standardized way, when the detailed categories are necessary.

Question 3.a - Is the example design seen in Figure 2 inclusive such that all individuals are represented?

The example design provides respondents with the ability to select or write in detailed race and ethnicity information, providing much flexibility. However, it is not fully inclusive, as some respondents might not know additional details about their racial or ethnic origins or might not feel comfortable disclosing this information. As highlighted in a 2020 article, the default categories primarily refer to countries of origin, which may not resonate with all respondents. For example, the article cited above describes how people of Ashkenazi Jewish descent have been unsure about how to respond to the detailed question. Similarly, people with Latin American roots, especially in Portuguese-speaking countries, might not see Hispanic or Latino as relevant to them. Adding criteria, would help respondents choosing Hispanic or Latino make a more meaningful choice.

To make this question more inclusive, the request to *Provide details below* should clearly indicate that it is optional or include an "I don't know" response.

Question 3.b - The example design seen in Figure 2 collects additional detail primarily by country of origin. What other potential types of detail would create useful data or help respondents to identify themselves?

Because race and ethnicity are multidimensional constructs, additional information might help respondents identify themselves more fully and therefore yield a more valid measurement system, as suggested by Viano and Baker (2020). These authors note that collecting language and recency of immigration in addition to country of origin might provide additional nuance. In addition, country of origin alone is unlikely to capture, and therefore help address, structural inequities within the broader federally recognized racial and ethnic groups. Similar to our recommendation under the response to Question 2a, we propose that the examples given should also seek to include ethnicities

that transcend countries (such as regional groupings), as questions that list examples can bias respondents to the options provided (Chestnut et al. 2007).

Question 3.c - Some Federal information collections are able to use open-ended write-in fields to collect detailed racial and ethnic responses, while some collections must use a residual closed-ended category (e.g., "Another Asian Group"). What are the impacts of using a closed-ended category without collecting further detail through open-ended written responses?

Using a residual closed-ended category without collecting further detail through open-ended responses would limit the ability to measure and address inequities. For example, a residual closed-ended category, such as "Another Asian Group," would mask inequities within still sizable ethnic groups, such as Cambodians, Laotians, and Hmong, who face different barriers than the other Asian groups listed in the response options (Southeast Asia Resource Action Center 2013). At the same time, open-ended written responses are more prone to response error, instability over time, and missing data (for example, see Viano and Baker 2020 and Aspinall 2009). In addition, collecting and coding open-ended responses increases the burden on respondents and agencies. One consideration might be to create residual groups that allow for some differentiation—for example, "Another Asian Group (Southeast Asia)," "Another Asian Group (East Asia)," and so on—to help curtail data loss.

Requiring respondents to choose from a lengthy list can be burdensome and mode variation may arise. While advances in electronic data collection can simplify the process, such as Kaiser Permanente's use of software that recognizes keystrokes to present only the most relevant options among their list of 250 granular ethnicity categories (Agency for Healthcare Research and Quality 2018), concerns about accuracy and comparability of granular race and ethnicity collected in other modes must be considered.

Question 3.d - What should agencies consider when weighing the benefits and burdens of collecting or providing more granular data than the minimum categories?

To fully realize the benefits of collecting more granular data, it will be important to provide agencies with guidance on how to effectively collect, analyze, and report data at this level. One of the primary concerns of OMB is the potential burden placed on the public, particularly marginalized or vulnerable populations, when participating in federal studies. The collection of more granular data than the minimum categories will increase the burden on respondents, and in studies with small sample sizes, quantitative analysis may not be feasible. However, in certain cases, collecting more detailed data may be warranted when studying and supporting specific marginalized communities that have been historically overlooked because of their small size. Guidance on balancing the need for collecting comprehensive data against the specific data needs, potential burden, sample sizes, and the impact on representativeness of small subgroup populations can provide standardization, transparency, and reproducibility to ensure appropriate use.

Given the real burdens of collecting and processing granular data, it is important for OMB to provide guidance on:

- *Use of granular data, such as methodological approaches for analyzing small sample sizes.* As one example, Forrow et al. (2023) used Bayesian hierarchical modeling to improve the reliability of subgroup measures for very small groups. The guidance should also address common barriers to using granular data. Typically, small groups are aggregated to larger categories because of concerns

about small sample sizes. However, this approach can obscure important disparities within groups (Schwabish and Feng 2021 and Maxim et al. 2023). Another barrier to expanding the use of more granular data is that findings for smaller groups are not reported consistently (or at all). Addressing these barriers and providing guidance on how to overcome them will be critical.

- *Standardizing data collection and reporting practices* such as those offered by the Agency for Healthcare Research and Quality (2018)—including which granular categories to consider, how to treat multiple granular responses, and how to incorporate granular responses to the larger categories when relevant..

- *Data collection mode when collecting granular data.* Multi-mode surveys, the best approach to increasing survey response rates, can suffer from mode bias, as different modes of data collection can elicit varying responses. Paper forms and web-based surveys allow respondents to visually see all response categories, but lengthy questions with multiple responses can be confusing and may lead to misinterpretation. Telephone surveys can present a cognitive burden as respondents must rely on auditory language to comprehend a lengthy set of response options presented by the interviewer. This can potentially impact the cognitive steps involved in responding. In addition, web-based surveys may present device effects, which should be considered for the growing percentage of respondents who use handheld devices to complete web surveys. Agencies must carefully assess and test the functionality and view of the longer item format to ensure compatibility with all electronic formats, including mobile phones. Paper forms with limited space will require careful consideration on how to present and fit the long item format.

Question 3.e - Is it appropriate for agencies to collect detailed data even though those data may not be published or may require combining multiple years of data due to small sample sizes?

It is important for agencies to assess their study goals and research questions when deciding on the level of detail needed when collecting data, and this also applies to the race and ethnicity questions. Collecting detailed data can be burdensome for respondents and OMB should assess when it is appropriate for agencies to administer the detailed questions. Providing options that allow for more explicit racial and ethnic identifications improves the accuracy of data collection and reflects diversity within certain OMB categories. For example, with the decline of stigma felt by Indigenous peoples to identify openly as Indigenous, the number of people self-reporting that they are America Indian or Alaska Native has increased over time, far faster than would be possible from natural population increase. We support the collection of these data when they are necessary to assess and address disparities within larger racial and ethnic groups. It is important for agencies to strike a balance between the level of detail necessary for their research goals and minimizing burden on respondents.

Question 3.f - What guidance should be included in SPD 15 or elsewhere to help agencies identify different collection and tabulation options for more disaggregated data than the minimum categories? Should the standards establish a preferred approach to collecting additional detail within the minimum categories, or encourage agencies to collect additional information while granting flexibility as to the kind of information and level of detail?

We recommend offering guidance on a preferred approach that enables local adaptation, allowing agencies to select from a list of standardized categories for collection and reporting. Standardization of minimum and detailed categories is critical, as different reporting standards across agencies limit the usability of the data collected (Gonzalez et al. 2022). In line with the recommendations offered

by the [Agency for Healthcare Research and Quality (2018)](#), we suggest developing a national, comprehensive standard set of categories that local agencies can then choose from based on the populations most relevant to their context; research shows that the most relevant detailed categories vary based on geography, urbanicity, and level of population diversity at the local level. The agency's recommended [category lists and coding schemes](#) offer a useful starting point. This approach would standardize the kind of information and level of detail collected and reported while still allowing for some flexibility.

Question 3.g - Is the current "default" structure of the recommendation appropriate? Should SPD-15 pursue a more voluntary approach to the collection of disaggregated data, as opposed to having a default of collecting such data unless certain conditions are met?

As noted in our response to Question 3.e, it is important to consider the goals of each data collection to determine whether a voluntary approach may be more appropriate in some circumstances rather than as an absolute requirement. While detailed data can provide valuable insights for specific subgroups, its length and complexity may contribute to an increase in missing data. However, for particular research studies, including the detailed data may be essential to understanding and addressing disparities. For example, a [2007 study of health disparities](#) found differences among ethnic subgroups even after controlling for other background characteristics.

Allowing individual respondents to opt out of providing detailed information and communicating clearly to respondents that the question is optional may be necessary. The decision to collect the disaggregated data should be driven by the goals of the study and use of the data.

Question 3.h - What techniques are recommended for collecting or providing detailed race and ethnicity data for categories with smaller population sizes within the U.S.

We recommend OMB consider [PolicyLink's guidance on disaggregating data on race and ethnicity to advance a culture of health.](#) This approach seeks to increase sample sizes of groups with smaller population sizes within the United States through oversampling, more extensive outreach, and offering additional options for in-language interviewing within these groups. OMB should provide guidance for additional detailed ethnicity categories that could be added as closed-ended responses in specific geographical locations with large populations of racial and ethnic groups that have a small overall population size within the United States. For example, OMB could recommend use of an Ethiopian response option in Montgomery County, Maryland, and an Afghan response option in parts of California.

We also recommend that OMB encourage data collection agencies to engage underrepresented communities to ensure its outreach efforts are appropriate and responsive to their needs and interests, and to address historical harms caused by misuse of data. This process requires not just gathering feedback from communities but restoring them as data experts through engagement that builds trust and minimizes harm ([Gonzalez at al. 2022](#)). In the health sector, [research suggests that assuring respondents that the collection of race and ethnicity data can improve care and resources for all can increase comfort level in reporting improve their own race and ethnicity](#), which can help reduce missing data. OMB could encourage agencies to including an explanation of the benefits of collecting race and ethnicity data, similar to the approach used in [a Commonwealth Fund study published in *the Journal of General Internal Medicine*](#), which found clear variations in respondents' comfort level when they were asked about race and ethnicity:

OMB could publish its own guidance or point to existing resources, such as the Engaging People with Lived Experience Toolkit or community-specific resources, such as guidance published by SAMHSA's Native American Center for Excellence.

To aid meaningful reporting, OMB should recommend methodological approaches for presenting information on racial and ethnic groups with smaller population sizes. We recommend OMB encourage the use of Bayesian hierarchical modeling to improve the reliability of data analyses based on small sample sizes (Forrow et al. 2023). However, OMB could consider alternative methods—for example, using qualitative information, such as testimonies, to define the significance of a quantitative finding based on a small sample, instead of basing significance on frequentist calculation of Type I error rate, per the Critical Race Quantitative Intersections + Testimonios approach summarized in Viano et al. (2020).

Question 4: Update Terminology in SPD 15

Question 4.a - What term (such as "transnational") should be used to describe people who identify with groups that cross national borders (*e.g.,* "Bantu," "Hmong," or "Roma")?

There is currently no evidence-based consensus on the best term to use when describing people that identify with groups that cross national borders. The choice of terminology may depend on the context and additional community-based research can help inform appropriate terminology.

1. *If a combined race and ethnicity question is implemented, what term should be used for respondents who select more than one category? For example, is the preferred term "multiracial," "multiethnic," or something else?*

   As implicit in the Working Group's proposed question stem for the combined question ("What is your race or ethnicity?"), these terms have different meanings and are not mutually exclusive. One refers to people with multiple racial backgrounds and the other refers to people with multiple ethnic backgrounds. One consideration is to simply refer to those who select more than one category as (for example) "respondents who select more than one ethnicity or race." We recommend conducting community-based research with a variety of populations to collect evidence on which terms should be used under specific contexts. Based on this evidence, OMB could develop guidance for agencies, promoting standardized decision-making and practices regarding the use of these terms.

2. *Please refer to Section D, Previously Tested Definitions of Minimum Categories. Are these draft definitions:*

   i. *Comprehensive in coverage of all racial and ethnic identities within the U.S.?*

   While the definitions attempt to provide examples of different identities, they may miss many complex and important nuances of race and ethnicity to be considered fully comprehensive. For example, the MENA definition misses large regions, such as the Gulf region, which is home to a significant population of Yemeni refugees living in the U.S. Race and ethnicity are social constructs based on physical characteristics, cultural expression, geographic regions, customs, histories, languages, and religions. Given the complexity of these constructs, developing a fully comprehensive list of all racial and ethnic identities that exist within the U.S. is likely an impossible task and may not be necessary. It is unclear how these definitions (which include details beyond what is offered in the proposed combined questions) will be used since they are not shown to respondents. While the Working Group should consider additional research to develop a more comprehensive range of racial and ethnic identities and provide clear guidelines

on how the definitions will be used going forward, it will also be important to assess respondent burden, if the definitions are to be used with respondents.

ii. *Using equivalent criteria?*

Without further information on how these definitions were derived, it is difficult to determine if they reflect equivalent criteria.

iii. *Reflective of meaningful distinctions?*

The draft definitions are useful in helping to make sense of the detailed categories and how the broader categories can be distinguished from one another.

iv. *Easy to understand?*

Some categories lack sufficient explanation, as noted in our response to Question 4.a2.i. For example, the MENA definitions, without context, do not capture the differences of self-identifying based on geographic location, language, or another distinction.

v. *Respectful of how people refer to themselves?*

Seeking feedback from the voices of the communities impacted by the use of these terms is the best way to understand whether a term is considered respectful, offensive, or out of date. Due to the ever-evolving nature of language, identity, migration, and social constructs of race and ethnicity, developing a plan that seeks continuous feedback both qualitatively and quantitatively from each community will be important.

3. *Please suggest any alternative language that you feel would improve the definitions.*

To improve the definitions, it must be acknowledged that responses can be affected by various factors, including the context and environment in which the question is asked, the wording and structure of the question, the identity of the person or agency asking the question, and the level of comfort and safety felt by the respondent. Gaining community-based input through focus groups, cognitive interviews, and pilot testing will help to test the definitions and capture nuances that will ultimately lead to more inclusive definitions.

Question 4.b - As seen in *Figure 2,* **based on the Working Group's initial proposal, the question stem asks "What is your race or ethnicity?" Do you prefer a different question stem such as: "What is your race and/or ethnicity?", "What is your race/ethnicity?", "How do you identify?", etc.? If so, please explain.**

We support the proposed approach of using the question "What is your race or ethnicity?" instead of combining the two terms with a slash. Using a slash may be viewed as insensitive to the diverse cultural experiences that people of different racial and ethnic backgrounds have. In addition, maintaining consistency in language around this question is also critical for accurate and reliable data collection across all modes. This is especially important for telephone interviews, as the interviewer needs to read the question stem verbatim to the respondents. Using a "slash" could lead to respondent confusion or agencies developing variations to overcome how the question is read over the phone, resulting in potentially inconsistent or inaccurate responses.

Question 5: Guidance is necessary to implement SPD 15 revisions on Federal information collections

Combining race and ethnicity into a single question and instructing respondents to select all that apply will likely increase the number of respondents who identify as multiracial. Current American Community Survey data shows that younger U.S. residents are more likely to report their ethnicity as multiracial, as shown in Figure 1. The data in the figure is based on the current two-question framework and does not include those identifying as Hispanic or Latino in the multiracial tabulations.

Figure 1. Share of each age cohort reporting two or more racial identity selections in the American Community Survey
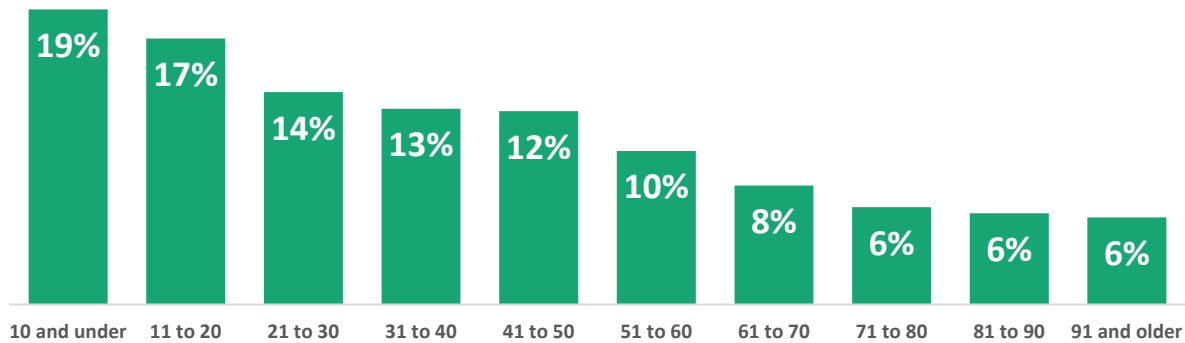
Source: ACS 1-Year Estimates Public Use Microdata Sample 2021

Question 5.a - For data providers who collect race and ethnicity data that is then sent to a Federal agency, are there additional guidance needs that have not been addressed in the initial proposals?

Providing guidance on how to combine or isolate categories based on sample sizes or analytic goals would be beneficial to researchers and analysts in better understanding and interpreting data. This is especially relevant for reporting multiracial and multiethnic identities as agencies implement Special Policy Directive 15. Aggregating individuals into a single "multiracial" category can result in a group with little in common, leading to difficulty in studying them and possible loss of meaning. Ultimately, each research or analysis question demands its own specific framework for aggregating race and ethnicity questions. For instance, a Mathematica evaluation of the Comprehensive Primary Care Plus Model, for the U.S. Department of Health and Human Services and the Centers for Medicare & Medicaid Services, identified a large proportion of beneficiaries who identified as American Indian or Alaska Native (AI/AN) and as one or more additional races; in this context, aggregating multiracial beneficiaries into a separate category of "multiracial" separate from beneficiaries who are American Indian or Alaska Native "only" drastically impacts the sample size of the AI/AN group. To fully understand the impact of the program on AI/AN beneficiaries, it is crucial to construct a subgroup of all AI/AN individuals, including those reporting multiple identities. Similarly, including Hispanic or Latino in the race categories might mean an uptick in people identifying as multiracial Hispanic. Therefore, guidance on knowing when to separate all Hispanic-identifying people for certain questions or reports will be crucial and help with standardization.

Moving away from using an undifferentiated "multiracial" category toward a more valid understanding of the complex and overlapping racial and ethnic identities of those in the United States might have the unintended outcome of producing myriad specific identities as people indicate belonging to two or more racial and ethnic groups. Depending on the population, these groups might be too small to analyze with sufficient statistical power. As a result, selecting more than one race or ethnicity might mean your data are thrown out of subgroup analyses. Therefore, guidance should recommend how to combine groups appropriately and sensitively to ensure analyses include all people. This guidance should not be prescriptive, as each study's estimation goals will determine whether and how groups should be combined, and whether they are multiracial or single but small race groups. Rather, the guidance should suggest how to think about combining groups given the study's goals.

Question 5.b - With the proposals to use a combined race and ethnicity question and to add MENA as a minimum category, what specific bridging concerns do Federal data users have? Please submit any research on bridging techniques that may be helpful to the Working Group. *Bridging refers to making data collected using one set of categories (e.g., two questions without MENA), consistent with data collected using a different set of categories (e.g., one question with MENA).*

We are not aware of any bridging techniques that could be useful for the specific bridging of the new MENA category as there is not an existing framework for this particular mapping. The introduction of this new category will disrupt trend data, but that should not deter the addition of this important new category. There is a long and rich history of the federal government changing how race and ethnicity data have been collected on the Census, indicating the constantly evolving constructs around these concepts and how they should be measured. The Working Group could look to how bridging has been done in the recent past as new and revised race and ethnic categories have been bridged to prior measures. Going forward with the new MENA category will allow for new trend data that is more accurate and inclusive.

Question 5.c - What guidance on bridging should be provided for agencies to implement potential revisions to SPD 15?

Guidance should clearly explain the appropriateness of prediction and imputation models for assigning race and ethnicity in limited situations when a respondent does not provide this information. It should also explain how to update models to use more complex multiracial identification that will result from changes according to Social Policy Directive No. 15. When a survey does not ask respondents to report race and ethnicity (or respondents or a knowledgeable proxy are asked for this information but do not report it), this information might be available from administrative data. However, all data collections should indicate the source of the race and ethnicity information. It is important to note that ensuring the accuracy of information that is not reported directly by the sample member is challenging.

Question 5.d - How should race and ethnicity be collected when some method other than respondent self-identification is necessary (*e.g.,*** by proxy or observation)?**

In cases where self-identification of race and ethnicity is not collected or cannot be linked to existing administrative data, it is important to consider whether race and ethnicity measures are necessary before determining a data collection approach. If dealing with item-level missing data, a common technique is imputation, which would require other variables in the data set to predict the missing

values. In cases where a proxy is already part of the study design, for example when an individual is completing the race and ethnicity question for all members of a household in the Census, the accuracy of the information obtained from the proxy may depend on the relationship of the proxy to various household members. Use of observations, as discussed in our response to Question 5.e are generally not recommended.

One approach that requires further research is use of predictive models. For example, although Medicare claims data do not contain self-reported racial or ethnic identity, the RAND Corporation and RTI International produce measures that predict race and ethnicity based on other information in the Medicare database. However, the use of predicted race and ethnicity raises a host of methodological, practical, and ethical questions. Important information to know when using predicted race and ethnicity includes (1) the information used to predict race and ethnicity (for example, residential address and surname information combined with existing race and ethnicity administrative data), (2) how the predicted measure defines the race and ethnicity categories and whether that differs from how the self-report categories are defined, and (3) how well race and ethnicity predictions correspond with self-reported identities (that is, the correlations between the predicted race and ethnicity and existing measures of self-reported race and ethnicity as a validation check). The use of predictive models requires more research to assess accuracy, explore key measures necessary for the model, and when use of a model may or may not be appropriate.

Question 5.e - What guidance should be provided for the collection and reporting of race and ethnicity data in situations where self-identification is unavailable?

In situations where self-identification is unavailable, it is important to recognize that attempting to guess someone's race or ethnicity based on their physical appearance can be inaccurate and demonstrates a lack of cultural humility. However, there are other resources that can be used in certain circumstances, although they will not be as reliable as self-identification. One option is to link to existing data source such as administrative data. Alternatively, proxy respondents, such as parents responding on behalf of their children, have been used in some cases for reporting of race and ethnicity data. Statistical imputation can be used to estimate item-level missing data. If possible, obtaining community-based feedback to ensure that the chosen strategy is appropriate and aligns with the cultural norms would be ideal, and should be guidance offered by OMB.

Question 6: Comments on Any Additional Topics and Future Research

Question 6.a - SPD 15 does not dictate the order in which the minimum categories should be displayed on Federal information collections. Agencies generally order alphabetically or by population size; however, both approaches have received criticism. What order, alphabetical, or by population size, do you prefer and why? Or what alternative approach would you recommend?

Currently there is no established, evidence-based consensus on how to order response categories when collecting data on race and ethnicity, and further research is necessary. While some federal guidance proposes listing the minimum categories alphabetically, the 2020 Census ordered the race categories primarily by population size. It will be important for OMB to provide guidance to agencies that minimize bias and allow for consistency in how this information is gathered and reported overtime.

Question 6.b - The current minimum categories are termed:
- **American Indian or Alaska Native**
- **Asian**
- **Black or African American**
- **Hispanic or Latino**
- **Native Hawaiian or Other Pacific Islander**
- **White**

Do you have suggestions for different terms for any of these categories?

Although widely accepted, the current terminology for the American Indian or Alaska Native category may not be fully inclusive of Indigenous people from other parts of North America, such as Mexico, Canada, and the Pacific Islands. Some Indigenous persons and their descendants from these areas may identify more readily with the term "Indigenous." However, there is currently no consensus on terminology for this diverse population. The Native American Journalists Association recommends using the term "Indigenous" as an identity rather than an adjective, as tribal membership or citizenship denotes identity. Further community-based research is necessary to determine which terminology is preferred and appropriate.

Regarding the terminology used for U.S. residents with family origins in Spanish-speaking countries and people who are of Latin American descent, there is disagreement about the use of gender-neutral alternatives such as "Latinx." Historically, Hispanic has referred to people from Spain and other Spanish-speaking countries, including Latin American countries, while Latino has described people from Latin American countries only. However, while "Latino" may be considered a gendered term, Latinx is not a term that is familiar to many in the Hispanic or Latino community. As such, it may be most appropriate to continue using Hispanic or Latino until further community-based research can provide opportunities for these groups to express how they wish to be referred to.

Question 6.c - How can Federal surveys or forms collect data related to descent from enslaved peoples originally from the African continent? For example, when collecting and coding responses, what term best describes this population group (e.g., is the preferred term "American Descendants of Slavery," "American Freedmen," or something else)? How should this group be defined? Should it be collected as a detailed group within the "Black or African American" minimum category, or through a separate question or other approach?

Additional research is necessary to develop a standardized approach for use of terms and questions related to the population of descendants of enslaved peoples in Federal surveys or forms. We recommend using community-based research methods to co-create the most appropriate and respectful term and question(s) for this population. This methodology involves working collaboratively with the community, acknowledging their expertise and lived experiences, to develop the best terms and question(s) including appropriate language, context, and explanation about the need for collecting this information.

It is important to weigh the burden on respondents against the need for meaningful data and intended analytic use. Requiring detailed questions have a downside if the data collected will not be used, either due to small sample sizes or because the analytic goals do not align, increasing burden on respondents, and potentially creating opportunities for unintended misuse of data. Therefore, we also recommend using community-based methods to inform standardized guidance that OMB can provide on *when* to collect this level of information to minimize burden on respondents.

Mathematica
Progress Together

mathematica.org   website