



April 7, 2023

Robert Sivinski, Ph.D.
Chair, Federal Interagency Technical Working Group
on Race and Ethnicity Standards
Statistical and Science Policy Branch
Office of Information and Regulatory Affairs
Office of Management and Budget

Dear Dr. Sivinski,

The American Statistical Association (ASA) appreciates this opportunity to respond to OMB's Request for Comment on the Initial Proposals For Updating OMB's Race and Ethnicity Statistical Standards, as invited in the [Federal Register of January 27th, 2023 \(88 FR 5375\)](#).

We applaud the Federal Interagency Technical Working Group on Race and Ethnicity Standards (Working Group) and its thorough, carefully considered initial proposals. Our comments are listed according to the format of the *Federal Register* notice. Questions regarding these comments can be directed to ASA Director of Science Policy, Steve Pierson, spierson@amstat.org.

1. Collect race and ethnicity information using one combined question

We strongly agree with the proposal in principle. Indeed, we are unaware of a universally agreed-upon distinction between race and ethnicity, and we are concerned that collecting data on "race" separately from data on "ethnicity" implicitly sends the message that race is a purely biologically determined category, an idea that is intertwined with the history of racist violence and oppression in the U.S [see e.g. Roberts, D. *Fatal Invention: How Science, Politics, and Big Business Re-Crete Race in the Twenty-First Century* (The New Press, 2011)].

Further, we appreciate the substantial testing completed by the Census Bureau demonstrating the reduction in the use of the reporting category "some other race" among persons identifying as Hispanic/Latino when a combined format is used. Additionally, and most importantly, we appreciate the cognitive testing by the Census Bureau, which indicated a strong preference by Hispanic/Latino persons to identify as a race, rather than an ethnicity.

To develop responsible implementation of the proposed change, however, specific and complementary additional testing beyond that conducted by the Census Bureau is needed. In

particular, we recommend testing of the combined format without the "some other race" option, as to our knowledge, testing of the proposed measurement change has not been done without the collection category of "some other race." (Only the U.S. Census Bureau offers respondents a "some other race" collection category; this is required by law.) The magnitude of this change, which affects all federal agencies collecting and reporting data on race/ethnicity, and the complexities in using state records to produce federal data warrant further testing to better identify– and responsibly plan for– differences in collection and reporting for the vast majority of federal agencies, which do not include a "some other race" category. Understanding the impact of the change in terms of cost, timeliness, and data quality benefit (including that necessary for bridging studies) for these agencies will require testing the combined format without the "some other race" option. We recommend this testing be conducted as soon as possible to inform the wording of any revisions to the standard and better prepare implementation guidance as needed.

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

We endorse in principle the creation of a new, separate minimum reporting category for MENA. The *Federal Register* Notice defines MENA as follows: "The category 'Middle Eastern or North African' includes all individuals who identify with one or more nationalities or ethnic groups with origins in the Middle East and North Africa. Examples include, but are not limited to, Lebanese, Iranian, Egyptian, Syrian, Moroccan, and Israeli." We understand that OMB may revisit the clarity and scope of this definition pending review of the requested public comments.

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

We cautiously support in principle this approach. However, implementation guidance should be provided in the following areas to support data quality and confidentiality.

- For race/ethnicity groups with smaller population sizes within the U.S, we recommend that collecting (very) detailed race/ethnicity should be guided by the research question, and that reporting should be guided as well by robustness of the estimate and disclosure concerns. The latter consideration may be especially important given recent and continuing development of privacy-enhancing technologies. Studies focused on best practices for reporting detailed race/ethnicity data (including aggregating across years or other forms of data blending) are warranted to inform implementation guidance.
- A residual closed-ended category (i.e., without collecting further detail through open-ended written responses) could negatively impact racial/ethnic subgroups that are already underrepresented in information collections. This would make it difficult to assess whether these groups are disproportionately affected by certain phenomena and, accordingly, to create tailored approaches to reduce these disparities. To address this in part, we recommend the phrase "you may report more than one group" be bolded or otherwise emphasized.

- We strongly agree with the need for guidance in reporting detailed race/ethnicity data (in addition to guidance for collection). However, there is some concern that permitting a great deal of flexibility in reporting will have unintended consequences. Reporting detailed categories likely will require more staff effort than reporting broad categories, given necessary reviews of statistical precision and disclosure avoidance. Any flexibilities in reporting should guard against a possible tendency to report using the broadest categories permitted by default. There is also a concern that without clear guidance, there will be so much variation in the level of detail reported that statistical comparisons will be more difficult and the data less useful.
- To this end, we recommend an exchange of agency practices, in the form of working papers, to describe current best practices in reporting more than one race/ethnicity. Such papers, perhaps disseminated through the Federal Committee on Statistical Methodology, could inform implementation guidance for Statistical Policy Directive No. 15.

Nonetheless, when weighing the benefits and burdens of collecting or providing more detailed data than the minimum categories, we hope that agencies will consider that collecting detailed race/ethnicity data by default is consistent with a future federal data infrastructure that promotes *data equity*, in part, through leveraging *blended data*, as described below.

- We note that historically, federal agencies have been reluctant to collect detailed race/ethnicity data because the resulting sample sizes would be too small for precision in estimates, could affect confidentiality protections, or both. However, in more recent visioning of the federal data infrastructure, there has been strong technical progress in both of these areas.
- A greater policy emphasis on *blending data*—either combining multiple rounds of a given data collection and/or combining with appropriate additional data—has led to federal policies and technical processes to support this work. At the same time, there has been growth in privacy-enhancing technologies that can be leveraged to protect confidentiality, particularly in blended data. This is being pursued in earnest both within the US and internationally.
- There has also been a complementary push for *data equity*, which considers, among other aspects, the utility of reported data to inform federal policy. (See for example, Executive Order 13985, [Advancing Racial Equity and Support for Underserved Communities Through the Federal Government](#) and [The Equitable Data Working Group Report](#)) If detailed race/ethnicity data are neither collected nor reported, then the experiences and needs of particular groups cannot be as well taken into account.
- However, it should be made clear that collecting such data does not require agency coding or analyzing write-in responses if resources at that agency do not so permit.

Coding and analysis of write-in responses could be undertaken by researchers for broader benefit.

4. Update Terminology in SPD 15

Assuming a combined race/ethnicity question is implemented, we strongly recommend the following:

- To avoid perpetuating the idea that race is a strictly biological category clearly different from ethnicity, **the question stem should not use the terms “race” and “ethnicity.”** One strong alternative is currently used by the *All of Us* project: "Which categories describe you? **Select all that apply.** Note, you may select more than one group."¹

However, to guard against the misperception that the question obligates respondents to indicate every aspect of their heritage, the exact language used should be informed by research. Then, reporting guidance should, to the fullest extent permissible given precision of the estimate and disclosure avoidance, indicate specific race/ethnicity for those who selected more than one (e.g., Hispanic (alone or with any other race), Hispanic and White, and Hispanic and Black).

If the words "race" and "ethnicity" are retained in the question stem, a (less preferred) recommendation is to include historical context to respondents and surveyors when collecting this information. However, we note that this adding such language may have the unintended effect of further politicizing the collection and measurement of race/ethnicity.

- A term such as “**multiracial**” or “**multiethnic**” should **not be a collection category or a reporting category.** It is not now a collection or reporting category. Data collected or reported in this category cannot be analyzed to inform federal policy.

5. Guidance necessary to implement SPD 15 revisions on Federal information collections

When respondent self-identification is not possible, we strongly recommend the reported data should indicate the source of the data. For example, we encourage the Working Group to distinguish between reporting by knowledgeable persons in the household (e.g., a person reporting their spouse's self-described race/ethnicity on behalf of their spouse) and reporting by "observation" (e.g., a school administrator). Although a person's perceived race may be highly relevant to their lived experience, it is a distinct datum from their self-described race.

¹<https://www.researchallofus.org/wp-content/themes/research-hub-wordpress-theme/media/2020/12/Basics.pdf>.

6. Comments on Any Additional Topics and Future Research

We again applaud the Working Group on taking up the review of Statistical Policy Directive 15. For reasons well-described in the *Federal Register* Notice requesting public comment, measurement of race and ethnicity is a fluid and continually changing concept. The technologies for collecting and reporting data also continue to evolve. At the same time, the salience of measurement remains constant.

To achieve their intended purposes, standards should not often change. Yet, implementation guidance, and even OMB guidance memoranda, are ways in which complementary updates to standards can be provided.

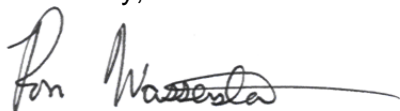
Similarly, managing a Federal Interagency Technical Working Group every several years to support systematic review and possible revision of a standard requires substantial time commitment. However, perhaps smaller groups, meeting more regularly and developing recommendations complementary to an existing standard for consideration by the Office of the Chief Statistician would require a lower time commitment and provide more timely feedback to support best federal measurement.

Accordingly, we encourage the Working Group to consider ways in which ongoing examination and guidance regarding measurement of race/ethnicity could support incremental improvement between comprehensive reviews of the standard. To that end, we note that Executive Order 13985 established an interagency Equitable Data Working Group (cited above), whose functions include identifying inadequacies in measurement of racial and other underserved communities in existing federal data-collection programs and facilitating the sharing of information and best practices, consistent with applicable law and privacy interests. We anticipate some opportunities to provide a venue for ongoing examination and development of race/ethnicity collections and reporting measures may thus become available.

In closing, I thank the ASA Committee on Professional Ethics and the following ASA members from various other ASA committees, groups, and sections for their contributions to these comments: Emily Butler, Sloka Iyengar, Kristian Lum, Stephanie Morales, Lauren Samuels, Joshua Snoke, Suzanne Thornton, and Ulrich Kemmo Tsafack.

Thank you again for the opportunity to comment on this important endeavor.

Sincerely,

A handwritten signature in black ink, appearing to read "Ron Wasserstein", with a long horizontal flourish extending to the right.

Ron Wasserstein
Executive Director