

**“Necessary Research for Revising OMB’s Race and Ethnicity Standards”**  
**By Yasmiyn Irizarry PhD, Edward Vargas PhD, Nancy López PhD**  
**April 26, 2023**

This comment summarizes our recommendations regarding necessary sources of information and avenues of further exploration for revising OMB standards for Federal Data on Race and Ethnicity. Importantly, researchers working outside of the U.S. Census Bureau, us three included, have, at best, limited access to the relevant information (and absolutely no access to the underlying data) used to develop and substantiate the initial Working Group proposals presented in this Federal Register. Accordingly, our initial recommendations focus on information and data sharing, which align with the White House Office of Science and Technology Policy (OSTP) current [policy guidance](#) (updated August 25, 2022) regarding public access to federally funded data.

#### INFORMATION & DATA SHARING

1. Provide a breakdown of costs for the preparation, implementation, and subsequent data cleaning, analysis, and report development associated with the 2010 Alternative Questionnaire Experiment (AQE) and the 2015 National Content Test (NCT).
2. Ensure that all forms (e.g., questionnaires, instrument specs) associated with the [NCT](#) are publicly available. Make AQE and NCT data accessible to researchers through Federal Statistical Research Data Centers. Create an expedited process for researchers proposals that aim to inform the revision process or discourse. This includes but is not limited to studies that (a) test Interagency Working group proposals or claims, (b) examine research findings from federal reports, (c) evaluate potential implications of proposed changes, or (d) address current debates or tensions.
3. Constitute a technical advisory committee comprised of scholars with expertise in race/ethnicity (alone or in combination with work in identity, intersectionality, inequality and/or civil rights/law) and quantitative methods or measurement to review existing and potential variables, questionnaires, and other documentation, make informed recommendations on avenues for further exploration, and evaluate the ability of a given categorization scheme and question format to illuminate inequalities in key outcomes (e.g., educational attainment, housing segregation, voting redistricting, incarceration, health and wellbeing). The committee should also include those who support the combined question as well as those who oppose. Appointed members should come from the communities represented by current and proposed OMB race and panethnic categories, and also reflect the racial and ethnic subgroup diversity within these communities.

## ISSUES RELATED TO EXISTING DATA

4. Revise the NCT report or produce a new report that meets the more rigorous reporting standards used in the AQE report. This report should include (a) sample sizes, (b) response rates, and (c) estimates of response bias and variance for the full NCT as well as by panel. These estimates should be reported for all race, ethnicity, and detailed origin groups, as well as for subgroups at the intersections of race and ethnicity/origin.
5. Using NCT data, estimate detailed frequency and percentage distributions of (a) OMB race and ethnicity categories and all of their detailed subgroups (rather than just the largest groups), (b) race within each panethnicity (i.e., Hispanic/Latinx; Middle Eastern and North African, American Indian and Alaskan Native), and (c) race within each detailed subgroup. Provide these distributions by race/ethnic group strata, by region (or state, if feasible), and by panel.
6. Use NCT data to test for differences in the reporting of race and ethnicity within each panel for Hispanic/Latinx respondents (altogether as well as by detailed ethnic groups and racial subgroups) sampled from different race/ethnicity group strata.
7. Use multivariate models to test whether differences in the percentage of Black alone, Black alone or in combination, or Afro-Latinx between panels in the NCT are related to or a result of differences in the detailed ethnic, strata, or state/regional composition of each panel.
8. Evaluate whether using perceived race to determine truth changes the interpretation of results in the NCT report, and if so, to what extent. Estimate detailed frequency and percentage distributions of perceived race for each OMB race and ethnicity category. Calculate distributions of the share of (dis)concordance between self-identified and perceived race within each panethnicity and detailed subgroup. Use multivariate techniques to examine the role of the respondent's race as perceived by others on (a) the reporting of race and ethnicity in each format and (b) shifts in the reporting of race and ethnicity when an alternative format is provided.
9. Combine data from the 2010 AQE and 2015 NCT to examine the impact of using the combined question format versus two separate questions for race and ethnicity on the counts for Black alone, Black alone or in combination, and AfroLatinx.
10. Identify who made the decision to include write-in space for origins under the White and Black/African American race boxes in the 2020 Census (e.g., this had never been done before as even the American Community Survey has a separate ancestry question). Investigate whether any testing was completed to assess the impact of this change on the identification, coding, reporting, and/or tabulation of race and ethnicity.
11. Make clear what share of a given estimate or pattern of change is due to (a) autocoding/

backcoding/ data editing based on OMB standards or (b) autocoding/ backcoding/ data editing based on Census (not OMB) standards (vs. respondent selection). This breakdown should be incorporated into the AQE and NCT reports. This information should also be provided for any other estimates, tables, and figures already created by the Census in support of this revision process, as well as for anything new they produce moving forward.

12. Reproduce analyses in the AQE and NCT reports using an alternative version of each dataset that includes in-line edits or pre-edits for each race (and ethnicity) but no change to other race or ethnicity boxes. Compare the results produced using this alternative version of data to the original findings in each report. This method can also be used to evaluate the impact of autocoding/backcoding/data editing on recent findings for and shifts in race and ethnicity reporting in the Decennial census and American Community Survey.
13. Use AQE and NCT data to examine whether the impact of question format differs for self-identification of race and ethnicity compared to the identification of race and ethnicity for someone else in their household. Evaluate whether this pattern differs when reporting the race and ethnicity of a child versus that of an adult.

## FUTURE TESTING & ALTERNATIVE FORMATS

14. Simulate actual Census conditions as testing conditions (e.g., include an all internet postcards, with paper only mailed after multiple attempts and in-person non-response follow-up sample)
15. Include a purposeful oversample of Black Latinx strata in subsequent studies to allow for definitive assessment of the impact of the two-part question versus combined question.
16. Test how changes in wording format impact the reporting of race and ethnicity. This may include wording that:
  - a. Clarifies for the respondent the difference between race and ethnicity.
  - b. Explains how their selections would be reported; not just existence of categories but likely tabulations.
    - i. Would respondents select more than one race box or provide a discordant origin if they knew that they would be categorized as multiracial? Respondents with multiple responses are not included in any of the racial categories selected as they can only be counted once for total population counts.
    - ii. Would respondents report different or no detailed ethnicity information if they disagreed with the group they were counted in? Respondents who report two or more ethnic groups within a given race or ethnicity are counted with the smallest ethnic group; Knowing this, a respondent who wishes to be counted as Black American, Foundational Black, or ADOS, rather than as Haitian-the smaller ethnic group-may choose not to report their Haitian origins.
  - c. Asks, “What is your race or visual racial status?” and “What is your ethnicity or cultural background and heritage?” rather than asking “What is your race or ethnicity. Mark all that apply.” (Race is not the same as ethnicity or cultural background, tribal affiliation or enrollment. Ethnicity is not the same as race which has a visual component).
17. Test alternative racial categories that can measure the racial inequities faced by those individuals who are not racialized as Black, White or Asian, but rather as “Brown,” “Mixed” or “Mestizo.”
18. Test differences in the reporting of race and ethnicity when MENA is included as an ethnic category rather than as a co-equal racial category.

19. Test alternative format of the two-part question. For example:

**What is your ethnicity, cultural background, or heritage? *Select all that apply***

**HISPANIC OR LATINO**

Mexican or Mexican

Cuban

Dominican

American Puerto Rican

Salvadorian

Colombian

*Print, for example, Guatemalan, Ecuadorian*

**MIDDLE EASTERN OR NORTH AFRICAN**

Lebanese

Egyptian

Moroccan

Iranian

Syrian

Israeli

*Print, for example, Algerian, Iraqi*

**ASIAN**

Chinese

Vietnamese

Korean

Native Hawaiian

Samoan

Filipino

Japanese

Chamorro

Asian Indian

Other Asian—*Print, for example, Pakistani, Hmong*

Other Pacific Islander—*Print, for example, Tongan, Fijian*

**AMERICAN INDIAN OR ALASKAN NATIVE**—*Print name of enrolled or principal tribe*

**OTHER ETHNICITY OR NATIONALITY**—*Print, for example, Canadian, Ethiopian, German, etc.*

**What is your race or visual racial status?**

BLACK, AFRICAN AMERICAN, OR AFRODESCENDENT

INDIGENOUS, INDIGENA, OR NATIVE

MESTIZO, MIXED, OR BROWN

ASIAN

WHITE

20. Proposed definitions based on alternative two-part question.

**ETHNICITIES**

**HISPANIC OR LATINO:** This category includes all individuals who identify as Hispanic or Latino, OR who identify with one or more nationalities or ethnic groups from Latin America, including in the Caribbean, Central America, and South America, as well as other Spanish cultures and select or report their ethnicity under the Hispanic or Latino category.

**MIDDLE EASTERN OR NORTH AFRICAN:** This category includes all individuals who identify as Middle Eastern or North African, OR who identify with one or more nationalities or ethnic groups from the Middle East or North Africa and select or report their ethnicity under the Middle Eastern or North African category.

**ASIAN:** This category includes all individuals who identify as Asian, OR who identify with one or more nationalities or ethnic groups originating in East Asia, Southeast Asia, or the Indian subcontinent, and select or report their ethnicity under the Asian category.

**NATIVE HAWAIIAN OR PACIFIC ISLANDER:** This category includes all individuals who identify as Native Hawaiian or with one or more nationalities or ethnic groups originating in Hawaii, Guam, Samoa, or other Pacific Islands.

**AMERICAN INDIAN OR ALASKA NATIVE:** This category includes all individuals who identify with one or more ethnicities or cultures of the original peoples of North, Central, and South America AND who maintain tribal affiliation or community attachment. It includes individuals who identify as American Indian and Alaskan Native, OR who report responses such as Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, or Nome Eskimo Community under the American Indian or Alaskan Native category.

**RACES**

**BLACK, AFRICAN AMERICAN, OR AFRODESCENDENT:** This category includes all individuals who identify as and/or are perceived by others as having origins in any of the Black racial groups of Africa.

**INDIGENOUS, INDIGENA, OR NATIVE TO THE AMERICAS:** This category includes all individuals who identify as and/or are perceived by others as having origins with any of the Indigenous Peoples or Native Populations of North, Central, and South America.

**MESTIZO, MIXED, OR BROWN:** This category includes all individuals who identify as and/or are perceived by others as being of mixed-race ancestry (often white and indigenous) and/or having origins to one of the phenotypically brown populations from Latin America, the Middle East and North Africa, the Indian continent, and Southeast Asia and the Pacific

Islands.

ASIAN: This category includes all individuals who identify as and/or are perceived by others as having origins to any of the populations of East Asia.

WHITE: This category includes all individuals who report their race as White, OR who identify as and/or are perceived by others as primarily having origins to any of the White racial groups.