BOSTON UNIVERSITY

Boston University Center for Antiracist Research

April 25, 2023

Submitted via www.regulations.gov

Office of Information and Regulatory Affairs,
Office of Management and Budget,
725 17th St. NW
Washington, DC 20503

Re: 88 FR 5375, OMB Docket No. 2023-0001, Boston University Center for Antiracist Research Formal Comment regarding of Notice of Initial Proposals For Updating OMB's Race and Ethnicity Statistical Standards

The Honorable Richard L. Revesz, Administrator of the Office of Information and Regulatory Affairs:

The Boston University Center for Antiracist Research ("the Center") is a nonpartisan, nonprofit, university-based center that convenes researchers, advocates, and practitioners to find novel and practical ways to understand, explain, and solve seemingly intractable problems of racial inequity and injustice. We foster research-based policy innovation, data-driven educational and advocacy campaigns, and narrative-change initiatives in an effort to build an antiracist society that ensures equity and justice for all.

This Comment relates to the Center's ongoing research and policy analysis regarding ways to improve racial and ethnic data collection in the United States.[1] We learned of the Office of Management and Budget's ("OMB") interest in revising its 1997 Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity ("SPD 15")[2] after staff at the Center met with the U.S. Chief Statistician and other executive branch officials regarding our report *Toward Evidence-Based Antiracist Policymaking: Problems and Proposals for Better Racial Data Collection and Reporting*.[3]

---

[1] *See* Bos. Univ. Ctr. for Antiracist Rsch., Response to Request for Information; Equitable Data Engagement and Accountability, 87 Fed. Reg. 54269 (Sept. 2, 2022) [hereinafter, Bos. Univ. Ctr. for Antiracist Rsch., Response to Request for Information], www.bu.edu/antiracism-center/files/2022/10/2022.10.3_Request_for_Information_Final.pdf; NEDA A. KHOSHKHOO, AVIVA GEIGER SCHWARZ, LUISA GODINEZ PUIG, CAITLIN GLASS, GEOFFREY S. HOLTZMAN, ELAINE O. NSOESIE & JASMINE B. GONZALES ROSE, BOS. UNIV. CTR. FOR ANTIRACIST RSCH., TOWARD EVIDENCE-BASED ANTIRACIST POLICYMAKING: PROBLEMS AND PROPOSALS FOR BETTER RACIAL DATA COLLECTION AND REPORTING (2022), www.bu.edu/antiracism-center/policy/policy-reports/toward-evidence-based-antiracist-policymaking.

[2] Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity, 62 Fed. Reg. 58782, 58782–90 (Oct. 30, 1997).

[3] KHOSHKHOO ET AL., *supra* note 1; Karin Orvis, *Reviewing and Revising Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity - OMB*, THE WHITE HOUSE: OFF. OF MGMT. & BUDGET BLOG (June 15, 2022), www.whitehouse.gov/omb/briefing-room/2022/06/15/reviewing-and-revising-standards-for-maintaining-collecting-and-presenting-federal-data-on-race-and-ethnicity/.

Subsequently, in November 2022, the Center convened 29 national experts—scholars and community advocates—to consider ways that the racial categories currently used for data collection can be improved to better track racism. This Comment provides insights based on our research and engagement with scholars and advocates on this important topic.

Racial and ethnic data collection can uncover racial inequities, and thereby reveal the racist policies and practices that cause those inequities. For example, better racial and ethnic data in the public health context can help "ensure that the groups suffering the worst receive the most attention, treatment, and resources."[4] In legal proceedings, racial and ethnic data can help uncover the unconstitutional influence of racial bias, potentially providing grounds for relief.[5] This Administration has rightfully recognized that such data "allow for rigorous assessment of the extent to which government programs and policies yield consistently fair, just, and impartial treatment of all individuals, including those who have been historically underserved, marginalized, and adversely affected by persistent poverty and inequality."[6] This work is essential in light of a massive and growing body of research documenting deeply rooted systemic racial inequities.[7]

To better understand and track racism, the categories used for racial and ethnic data collection must closely approximate racialized experiences. Because responsible data collection requires respondents to self-identify, the categories must be recognizable to respondents and resonate with how respondents see themselves. The more these categories reflect racialized realities, the better we can understand how racism manifests and how it can be mitigated. Conversely, if the categories used for data collection conflate or are otherwise at odds with racialized experiences, then data collection will bring us no closer to the goal of understanding and ending racism.

---

[4] KHOSHKHOO ET AL., *supra* note 1, at 54.

[5] *See, e.g.*, Brief of Amici Curiae Boston University Center For Antiracist Research, Fred T. Korematsu Center for Law and Equality, Center on Race, Inequality, and the Law, and Criminal Justice Institute at Harvard Law School at 10–12, Commonwealth v. Mattis, Nos. 11693, 09265 (Mass. argued Feb. 6, 2023), https://www.bu.edu/antiracism-center/files/2023/01/Mattis-Amicus-Brief.pdf (arguing that racial disparities among 18-20-year-olds sentenced to life without parole in Massachusetts contributes to unconstitutional cruelty of those sentences).

[6] Request for Information; Equitable Data Engagement and Accountability, 87 Fed. Reg. 54269, 54269–70 (Sept. 2, 2022); *see also* Advancing Racial Equity and Support for Underserved Communities Through the Federal Government, Exec. Order No. 13985, 86 Fed. Reg. 7009 (2021), www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/; Further Advancing Racial Equity and Support for Underserved Communities Through The Federal Government, Exec. Order No.14091, 88 Fed. Reg. 10825 (2023), www.whitehouse.gov/briefing-room/presidential-actions/2023/02/16/executive-order-on-further-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/.

[7] *See, e.g.*, ARIANE HEGEWISCH & CHANDRA CHILDERS, INST. FOR WOMEN'S POL'Y RSCH., BLACK WOMEN TO REACH EQUAL PAY WITH WHITE MEN IN 2130 (2020), https://iwpr.org/iwpr-issues/esme/black-women-to-reach-equal-pay-with-white-white-men-in-2130/; Neda Maghbouleh, Ariela Schachter & René D. Flores, *Middle Eastern and North African Americans May Not Be Perceived, nor Perceive Themselves, to Be White*, 119 PROC. NAT'L ACAD. SCIS., no. 7, 2022, https://doi.org/10.1073/pnas.2117940119; Drishti Pillai, Nambi Ndugga & Samantha Artiga, *Health Care Disparities Among Asian, Native Hawaiian, and Other Pacific Islander (NHOPI) People*, KAISER FAM. FOUND. (May 27, 2022), www.kff.org/racial-equity-and-health-policy/issue-brief/health-care-disparities-among-asian-native-hawaiian-and-other-pacific-islander-nhopi-people/.

The Center submits this Comment in support of moving to a single question format for collecting racial and ethnic data (Section I); adding a "Middle Eastern or North African" category (Section II); requiring further disaggregation of data (Section III); amending the terminology used in the current standards (Section IV); and conducting additional research to further improve the categories for the purpose of tracking and addressing racism (Section V). While these changes will not result in a perfect model for data collection, they are steps in the right direction to track and mitigate racism. As discussed below, we recommend a regular review process to further improve SPD 15 and promote its alignment with evolving conceptions of race, which is continually constructed.

> **I.   OMB Should Move to a Single Question to Collect Racial and Ethnic Data**

The Center's support for a combined question about racial and ethnic identity is informed by our goal of tracking and, in turn, mitigating experiences of racism. A combined question that is carefully constructed can improve both responsiveness to the survey questions and respondents' ability to accurately reflect more aspects of their identity and experience. To be effective, however, a combined question must distinguish between distinct concepts of race on the one hand, and ethnicity and national origin on the other. Categories reflecting ethnicity or national origin should not be designated as set "subcategories" of racial categories. Rather, the question should allow for respondents to select any ethnicity or national origin category irrespective of their selection of a racial category, to make it clear that a person's ethnicity or national origin is not circumscribed by how they racially identify. While the combined question poses some challenges, data analytic techniques can—and should—be used to help address these challenges. Ultimately, the benefits of the combined question counsel for its adoption. Our recommendations are discussed further below.

> a. A Combined Question Increases Responsiveness, Which Improves the Collection of Information About Racism.

As an initial matter, research conducted by the U.S. Census Bureau indicates that when the standard two-question format for soliciting racial and ethnic data is replaced with a single, combined question, the number of people who provide complete responses to the question improves.[8]

Under the current two-part question, an increasing number of respondents are selecting "Some Other Race"—an indication that the current OMB categories are lacking and do not reflect respondents' racialized experiences.[9] In the 2020 Census, "[t]he Some Other Race population was the second-largest alone or in combination race group, comprising

---

[8] KELLY MATTHEWS, JESSICA PHELAN, NICHOLAS A. JONES, SARAH KONYA, RACHEL MARKS, BEVERLY M. PRATT, JULIA COOMBS & MICHAEL BENTLEY, U.S. CENSUS BUREAU, 2015 NATIONAL CONTENT TEST RACE AND ETHNICITY ANALYSIS REPORT 83-84 (2017), www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2015nct-race-ethnicity-analysis.pdf.

[9] KHOSHKHOO ET AL., *supra* note 1, at 56 (citing Hansi Lo Wang, *1 in 7 People Are 'Some Other Race' on the U.S. Census. That's a Big Data Problem*, NPR, www.npr.org/2021/09/30/1037352177/2020-census-results-by-race-some-other-latino-ethnicity-hispanic (last updated Oct. 2, 2021)).

15.1% of the total population."[10] Heavy reliance on the "other race" category can lead to data confusion by aggregating experiences of racism and obscuring the true extent of inequities.[11] Research indicates that moving to a combined question reduces selections of "Some Other Race," improving the accuracy and comprehensiveness of racial data collection efforts.[12]

Additionally, research indicates that under the current two-part question format, many respondents do not answer the question about race at all.[13] A combined question can improve responsiveness, increase consistency of responses over time, and cut administratively invalid selections in half, ensuring that a more complete picture can be drawn from the data.[14]

The data collection problems presented by the two-part question—that is, the selection of "Some Other Race" and nonresponse—contribute to the persistent underestimation and undercounting of the population that otherwise identifies as "Hispanic or Latino." [15] An estimated 94% of respondents selecting "Some Other Race" alone are people who identify as "'Mexican,' 'Latino,' and other Hispanic origin groups." [16] When the "Hispanic or Latino" population is misrepresented or obscured, it is impossible to evaluate how this population is impacted by racist policies or practices.[17] The combined question could mitigate this problem. Indeed, according to testing by the U.S. Census Bureau, "a significantly higher percentage of Hispanic reinterview respondents did not identify as Hispanic in the self-response survey when responding to the two-part question, compared to either of the combined question formats. It is not immediately clear why this

[10] *Id.* (citing Nicholas Jones, Rachel Marks, Roberto Ramirez & Merarys Ríos-Vargas, *2020 Census Illuminates Racial and Ethnic Composition of the Country*, U.S. CENSUS BUREAU (Aug. 12, 2021), www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html).

[11] *Id.*

[12] KELLY MATTHEWS, JESSICA PHELAN, NICHOLAS A. JONES, SARAH KONYA, RACHEL MARKS, BEVERLY M. PRATT, JULIA COOMBS & MICHAEL BENTLEY, U.S. CENSUS BUREAU, 2015 NATIONAL CONTENT TEST RACE AND ETHNICITY ANALYSIS REPORT 83 (2017), www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2015nct-race-ethnicity-analysis.pdf.

[13] *Id.* at 41.

[14] *Id.* at 41, 84–85.

[15] The Center recognizes that the terms "Hispanic" and "Latino" are controversial and debated. "Hispanic" has a colonial history. The term de-emphasizes Latino/a/e connection to the Americas and emphasizes Spanish heritage over Indigenous and African heritage. "Hispanic" also excludes the population descended from Latin America who do not share Spanish as a heritage language, but who may have similar racialized experiences in the United States. "Latino" can be perceived as lacking the gender inclusivity of other categories, which is concerning. At the same time, more gender-inclusive terms are not as recognized by much of the "Hispanic or Latino" population. As term recognition is paramount when discussing categories used for racial and ethnic data collection, we use "Hispanic or Latino" here with the awareness that they may be imperfect. Erin Blakemore, *'Hispanic'? 'Latino'? Here's Where the Terms Come from*, NAT'L GEOGRAPHIC (Feb. 10, 2022), www.nationalgeographic.com/history/article/hispanic-latino-heres-where-terms-come-from; Shannon Greenwood, *About One-in-Four U.S. Hispanics Have Heard of Latinx, but Just 3% Use It*, PEW RESEARCH CENTER'S HISPANIC TRENDS PROJECT (2020), www.pewresearch.org/hispanic/2020/08/11/about-one-in-four-u-s-hispanics-have-heard-of-latinx-but-just-3-use-it/ (last visited Mar 23, 2023).

[16] Eric Jensen, Nicholas Jones, Megan Rabe, Beverly Pratt, Lauren Medina, Kimberly Orozco and Lindsay Spell, *The Chance That Two People Chosen at Random Are of Different Race or Ethnicity Groups Has Increased Since 2010*, U.S. Census Bureau (Aug. 12, 2021), https://www.census.gov/library/stories/2021/08/2020-united-states-population-more-racially-ethnically-diverse-than-2010.html.

[17] Matthews et al., *supra* note 12 at 4, 45–48

would occur for the separate question, but the takeaway is that the combined question formats are achieving more consistent reporting among people who identify as Hispanic."[18] Moreover, some entities collect or analyze racial data using hybrid categories like Hispanic-White or Hispanic-Black, leaving no option to accurately reflect the experiences of people who identify as Hispanic or Latino, but neither White nor Black.[19]

Some scholars and advocates have raised concerns about the impact of a combined question on data collection regarding the racialized experiences of people who identify as Afro-Latino. These concerns include the reliability of the Census study results due to insufficient representation of the Afro-Latino population in the testing protocol, and inadequate engagement with the Afro-Latino community by the Working Group.[20] Additionally, scholars and advocates have raised the concern that adding "Hispanic or Latino" as a racial category inaccurately portrays the population who identifies as "Hispanic or Latino" as racially monolithic, when that category actually includes a diversity of racialized experiences.[21] The Center supports ongoing engagement with scholars and advocates who have raised these concerns, to ensure that any changes to the OMB categories are implemented in a manner that will not undercount Afro-Latino or other Latino populations.

Latinos are racially and ethnically diverse. It is important that Afro-Latino, Indigenous Latinos, White Hispanics, Asian Latinos, people who identify both racially and ethnically as Latino or Hispanic, as well as other Latinos, are offered a racial category that corresponds to their identity. We believe that our suggested question format in Figure A improves upon the working group's recommendation in this regard insofar as Figure A decouples the racial categories from the categories based on ethnicity and national origin. For example, as posed in Figure A, the question permits a respondent to identify as both "Black," and any category based on ethnicity or national origin, such as "Mexican" or "Cuban."[22] We also emphasize that OMB should provide clear guidance indicating that respondents can select more than one racial category, so that respondents know that they may self-identify as both "Black" and "Hispanic or Latino," for instance.

It appears that a large percentage of Latinos identify racially as "Hispanic or Latino" and not Black, White, Asian, or American Indian/Alaskan Native. The absence of "Hispanic or Latino" as a racial category has contributed to incomplete racial data collection. Adding a racial category that focuses on people who are not racialized as—or who do not identify as—White, Black, Asian, or American Indian/Alaska Native, but instead are

---

[18] *Id.* at 47.

[19] Vincent C. Allen, Jr., Christina Lachance, Britt Rios-Ellis & Kimberly A. Kaphingst, *Issues in the Assessment of 'Race' among Latinos: Implications for Research and Policy*, 33 HISP. J. BEHAV. SCIS. 411, 413 (2011). https://doi.org/10.1177/0739986311422880; Gene Demby, *On the Census, Who Checks 'Hispanic,' Who Checks 'White,' And Why*, NPR (June 16, 2014), https://www.npr.org/sections/codeswitch/2014/06/16/321819185/on-the-census-who-checks-hispanic-who-checks-white-and-why.

[20] NANCY LÓPEZ, POINTS OF CONSIDERATION: WHY COUNTING RACE FOR LATIN@S/HISPANICS IS IMPORTANT 1–2, https://static1.squarespace.com/static/63e409a2024c2f1c2e88c04f/t/63e55605d6bc366817095778/1675974149364/Latino+is+Not+a+Race+Fact+Sheet.pdf (last visited Mar. 22, 2023).

[21] *Id.*; *see also* TANYA KATERI HERNANDEZ, RACIAL INNOCENCE: UNMASKING LATINO ANTI-BLACK BIAS AND THE STRUGGLE FOR EQUALITY (2022).

[22] *See infra* Part I.b.

racialized or identify as "Hispanic or Latino" will improve the ability to track racism impacting Latinos. This is especially important because some legal frameworks and data reporting requirements currently focus only on racial data and not responses to the ethnicity question. As posed in Figure A, the combined question including "Hispanic or Latino" as a racial category enhances rather than diminishes the ability of Afro-Latinos, Indigenous Latinos, White Hispanics, Asian Latinos, and other Latinos to more accurately select their racial identities.

Given indications that a combined question improves overall reporting of the population that racially identifies as "Hispanic or Latino," and that the question proposed at Figure A would allow for more complex and nuanced data collection across racial and ethnic/national origin categories, we believe that a combined question can provide more accurate data about the racial, ethnic, and national origin diversity in Latino communities and the racism experienced by these communities.

> b. A Combined Question Must Distinguish Between Ethnicity/National Origin and the Construct of Race to Provide More Comprehensive Racial and Ethnic Data Collection.

To better track racism, a combined question must distinguish between the construct of race and ethnicity/national origin. The question should allow respondents to select the options that most closely reflect their experiences. Respondents' selections of ethnicity/national origin should not be limited based on their selection of a racial category, or vice versa. Below we have reproduced Figure 2, which reflects the combined question example shared by the Federal Interagency Technical Working Group on Race and Ethnicity Standards ("Working Group"). We have also generated Figure A, which reflects the Center's recommendation for a combined question that would allow for more accurate and comprehensive racial and ethnic/national origin data collection.[23]

The Working Group's proposed combined question, reflected in Figure 2, lists categories in a way that suggests a respondent can only select from a small set of ethnic/national origin categories based on their selection of a racial category. This erroneously implies that certain ethnic/national origin designations are subcategories that "belong" to associated racial categories. The combined question as posed in Figure 2 thus limits a respondent's ability to comprehensively self-identify and skews data collection. The question posed in this manner also improperly conflates the concept of race with ethnicity and national origin. **Race** is a position within a power construct that categorizes groups of people loosely based on limited phenotypical characteristics and (mis)perceptions of broad-stroke ancestral origination. **Ethnicity** is one's culture, traditions, or heritage often related to specific geographical region, language/linguistic usage, religion/faith/spirituality, customs, dress, or expression. **National Origin** refers to the nation or country where a person or their ancestors were born or originated.[24]

---

[23] We recognize that combining categories based on ethnicity and national origin also conflates two distinct concepts, and thus hinders data collection on ethnicity and national origin individually. The question presented in Figure A represents an administrable improvement on both the current question and the Working Group's proposed question in Figure 2, though we recommend further research and testing of subcategories that decouple ethnicity and national origin. *See infra* Part V.e.

[24] *DEI Toolkit: Ethnicity & National Origin*, AAUW, www.aauw.org/resources/member/governance-tools/dei-toolkit/dimensions-of-diversity/ethnicity-national-origin/ (last visited Mar. 22, 2023).

By contrast, the Center's proposed version of the combined question in Figure A distinguishes between race on the one hand and ethnicity or national origin on the other. This approach allows respondents to select categories based on ethnicity or national origin irrespective of race. Similarly, Figure 2 lists tribal affiliation as a subcategory of the American Indian and Alaska Native ("AIAN") racial category, which improperly conflates a political designation with race.[25] The Center's proposed version (Figure A) includes a fill-in box for tribal affiliation that is distinct from the categories reflecting race.

The Center's proposed version of the combined question (Figure A) also differs from the Working Group's proposed question (Figure 2) because it lists ethnic and national origin categories alphabetically, rather than by population size. Listing these categories by population size is suggestive due to selection bias towards categories that appear towards the top of the list, and can thus mask inequities.[26] While we understand the statistical basis for selecting ethnic and national origin identity groups by estimated population size,[27] the Center recommends against presenting ethnic and national origin categories in order of population size, as this could skew responses toward larger more familiar categories listed first and decrease responsiveness to categories listed last, even if respondents would otherwise identify with them. Listing the categories in alphabetical order will minimize selection bias and improve standardization over time, since population numbers will fluctuate but alphabetical ordering will remain constant.

---

[25] Brief of 497 Indian Tribes and 62 Tribal and Indian Organizations as *Amici Curiae* in Support of Federal and Tribal Defendants at 19, Haaland v. Brackeen, 142 S. Ct. 1205 (2022) (Nos. 21-376, 21-377, 21-378, 21-380), 2022 WL 3682220, www.supremecourt.gov/DocketPDF/21/21-376/234098/20220819163105655_21-376%2021-377%2021-378%2021-380ac497Tribesand62TribalOrganizations.pdfA.pdf ("[T]ribal membership decisions are decisions of tribal self-governance, not racial categorization.").

[26] Sean P. Mackinnon & Mengyao Wang, *Response-Order Effects for Self-report Questionnaires: Exploring the role of Overclaiming Accuracy and Bias*, 16 JOURNAL OF ARTICLES IN SUPPORT OF THE NULL HYPOTHESIS 113, 120-23 (2020).

[27] While our proposed combined question (Figure A) reflects the ethnic/national origin categories that the interagency working group identified based on population size, ultimately, we recommend further research on a more equity-based approach to selecting the ethnic/national origin categories that are listed on the question form. *See infra* Part V.e.

Figure 2: Federal Interagency Technical Working Group on Race and Ethnicity Standards 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

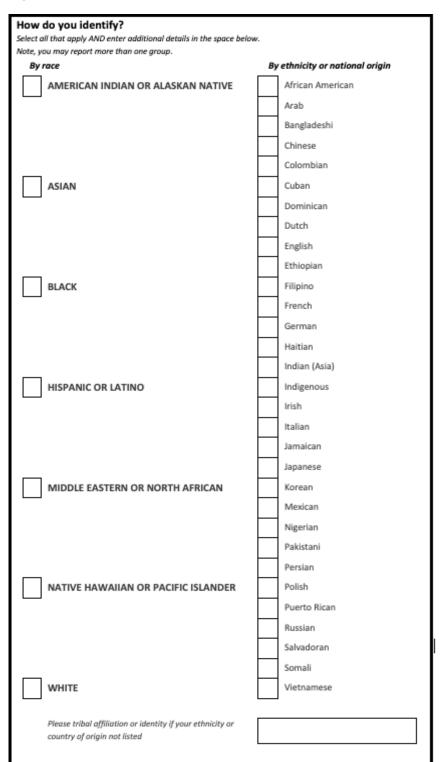*Enter, for example, Palauan, Tahitian, Chuukese, etc.*

Figure A: The Center's Recommendation for a Combined Question

How do you identify?

Select all that apply AND enter additional details in the space below.

Note, you may report more than one group.

By race	By ethnicity or national origin
☐ AMERICAN INDIAN OR ALASKAN NATIVE	☐ African American
	☐ Arab
	☐ Bangladeshi
	☐ Chinese
	☐ Colombian
☐ ASIAN	☐ Cuban
	☐ Dominican
	☐ Dutch
	☐ English
	☐ Ethiopian
☐ BLACK	☐ Filipino
	☐ French
	☐ German
	☐ Haitian
	☐ Indian (Asia)
☐ HISPANIC OR LATINO	☐ Indigenous
	☐ Irish
	☐ Italian
	☐ Jamaican
	☐ Japanese
☐ MIDDLE EASTERN OR NORTH AFRICAN	☐ Korean
	☐ Mexican
	☐ Nigerian
	☐ Pakistani
	☐ Persian
☐ NATIVE HAWAIIAN OR PACIFIC ISLANDER	☐ Polish
	☐ Puerto Rican
	☐ Russian
	☐ Salvadoran
	☐ Somali
☐ WHITE	☐ Vietnamese
Please tribal affiliation or identity if your ethnicity or country of origin not listed	[ ]

9

In sum, the use of a combined question with detailed ethnic and national origin categories would produce a "more accurate portrait of how the U.S. population self-identifies," especially for people who self-identify as multiracial or multiethnic.[28] More accurate and comprehensive data collection would, in turn, promote enforcement of civil rights laws and support antiracist policymaking.

We are mindful that the use of a combined question will require a careful approach to data analysis. Since the combined question would include seven minimum reporting category responses—rather than the current five minimum category responses to the question about race and one category response to the question about ethnicity—there is reason to expect the number of people selecting multiple categories will increase. This may amplify data challenges that already exist under the two-part question, wherein data from people who select more than one category may be improperly aggregated or erroneously analyzed. For example, Center research regarding the racialized impact of COVID-19 revealed that some states count populations selecting more than one racial category as two or more separate people.[29] Not only is this inaccurate, but research shows that double counting can lead to underestimation of racial inequities.[30] Additionally, the majority of states "reported COVID-19 data by reclassifying those who selected two or more races into a separate multiracial category, or in a combined category with 'Some Other Race.'"[31] This is problematic because aggregating everyone who selects more than one category risks obscuring distinct experiences of racism.[32] To address this issue, respondents who select more than one category should be included in any analysis of each category they select, respectively. For example, an analysis of respondents who identify as Black should include respondents who select "Black" alone *and* respondents who select "Black" in combination with any other category.

Finally, OMB should encourage further research regarding the possibility of collecting data not only about a respondent's identity but also their racialized experience—what some scholars have referred to as "street race."[33] A question such as this could further clarify the distinction between race and ethnicity, which are conflated in the Working Group's current proposed question form (Figure 2).

---

[28] Jones et al., *supra* note 10.

[29] KHOSHKHOO ET AL., *supra* note 1, at 26.

[30] *Id.* at 27–28.

[31] *Id.* at 26.

[32] *Id.* at 56 (noting that "[h]eavy reliance on the "other race" category can lead to data confusion and obscures the true extent of inequities").

[33] Nancy López, Edward Vargas, Melina Juarez, Lisa Cacari-Stone & Sonia Bettez, *What's Your "Street Race"? Leveraging Multidimensional Measures of Race and Intersectionality for Examining Physical and Mental Health Status among Latinxs*, 4 SOCIO. RACE & ETHNICITY 49 (2018), https://doi.org/10.1177/2332649217708798; DULCE GONZALEZ, NANCY LÓPEZ, MICHAEL KARPMAN, KARISHMA FURTADO, GENEVIEVE M. KENNEY, MARLA MCDANIEL & CLAIRE O'BRIEN, OBSERVING RACE AND ETHNICITY THROUGH A NEW LENS: AN EXPLORATORY ANALYSIS OF DIFFERENT APPROACHES TO MEASURING "STREET RACE" (2022), www.urban.org/research/publication/observing-race-and-ethnicity-through-new-lens.

II.   OMB Should Add a "Middle Eastern or North African" Category

There is a strong need for a Middle Eastern or North African ("MENA") category because the current standards completely obscure racism experienced by people who identify with this population. The Center supports including MENA among the racial categories, as the MENA population encompasses a broad and fluid set of ethnic groups who are often racialized as non-White based on perceptions of broad-stroke ancestral origination, and experience significant racism.[34]

During our convening on data equity in November 2022, scholars and advocates emphasized the urgent need to add the MENA category to understand how policies impact the MENA population. Many individuals who identify as part of the MENA population are generally not perceived to be White and do not perceive themselves as White.[35] If the MENA population is recorded as White in data collection, it is impossible to observe how policies impact this population. Potential policy impacts include housing segregation and discrimination, linguistic isolation, over-policing, environmental racism, healthcare disparities, and employment discrimination.[36] Moreover, when people do not feel that any of the listed racial categories reflect their race, their only option is to check "Some Other Race." As discussed above, an increasing number of people have chosen to select "Some Other Race" in recent years.[37] Indeed, in the 2020 Census, the "Some Other Race" population was the second-largest alone or in combination race group.[38] The broad reliance on the "Some Other Race" category reflects a systematic failure to provide appropriate racial categories and can lead to data confusion and improperly aggregates race, obscuring the extent of racial inequities.

There are questions about the appropriate terminology to use when describing this population. Some suggest that because the term Middle Eastern is a colonial designation,[39] the OMB should consider alternative terms like West Asian North African[40] or Southwest Asian North African ("SWANA").[41] However, scholars and advocates have noted that limited knowledge and use of these alternative terms increases the risk that respondents would not recognize or select this category.

It is important to consider the importance of term recognition while also acknowledging

---

[34] *See* Neda Maghbouleh, *From White to What? MENA and Iranian American Non-White Reflected Race*, 43 ETHNIC & RACIAL STUD. 613, 623 (2020), https://doi.org/10.1080/01419870.2019.1599130.

[35] Maghbouleh, Schachter & Flores, *supra* note 7; Khaled A. Beydoun, *Boxed In: Reclassification of Arab Americans on the U.S. Census as Progress or Peril?*, 47 LOY. U. CHI. L.J. 693, 695–702 (2016) (contextualizing OMB's role in the current classification of Arab Americans as White and explaining effect of a change in the census).

[36] *E.g.*, Daniel Widner & Stephen Chicoine, *It's All in the Name: Employment Discrimination Against Arab Americans*, 26 SOCIO. F. 806, 809, 818–19 (2011) (discussing results of study on employment discrimination against Arab Americans since 9/11); Maghbouleh, Schachter & Flores, *supra* note 7 (revealing ongoing disparities and inequalities faced by MENA Americans).

[37] Jones et al., *supra* note 10.

[38] *Id.*

[39] Yara M. Asi, *The Colonial Legacy in the Arab World: Health, Education, and Politics*, ARAB CTR. WASH. DC (Nov. 9, 2022), https://arabcenterdc.org/resource/the-colonial-legacy-in-the-arab-world-health-education-and-politics/.

[40] *Why WANA?*, WANA INST. wanainstitute.org/en/why-wana (last visited Mar. 22, 2023);

[41] Rayya El Zein, *Introduction: Cultural Constructions of Race and Racism in the Middle East and North Africa / Southwest Asia and North Africa*, LATERAL, Spring 2021, https://doi.org/10.25158/L10.1.11.

that terminology and associated perceptions change over time. Questions about the appropriate term are best resolved through a regular procedure to reevaluate and update SPD 15.[42] Such questions do not warrant the continued omission of the MENA category, given the strong interest in tracking and addressing racism experienced by this population.

III.   OMB Should Require Disaggregated Data

The current OMB categories are overly broad, impeding efforts to measure and advance equity.[43] The proposed revisions (Figure 2) suggest an improvement insofar as the addition of ethnic and national origin categories creates more opportunity for disaggregation, though as discussed above, we recommend decoupling those categories from the racial categories (Figure A).

In addition to the revisions OMB plans to make to SPD 15, the Center recommends amending the guidance to require—rather than merely encourage—federal agencies and other data collecting entities to further disaggregate categories that they use for collection based on the needs and constituencies of their respective jurisdictions. The Center's efforts to research racial and ethnic inequities have been limited by the amount of granularity reflected in current datasets.[44] A lack of granularity impedes analyses of the ethnic and racial makeup of jurisdictions and inequities between subpopulations. Agencies and localities should determine what subcategories to add to their data collection efforts, as needs may vary based on their constituencies. For example, one jurisdiction might disaggregate "Asian" into Southeast Asian and East Asian, while a particular agency might have reason to disaggregate "White" into groups like Eastern European and Western European.[45] These subcategories should be created such that they can be collapsed into OMB's minimum categories used for data collection, to ensure compatibility across datasets and preserve longitudinal data if categories change over time.

IV.   OMB Should Amend Specific Terminology and Definitions

a.   Remove the Terms "Negro, "Minority," and "Majority"

The term "Negro" is antiquated and broadly understood to denote inferiority of people who are racialized as Black.[46] The Center strongly supports the Working Group's

---

[42] Bos. Univ. Ctr. for Antiracist Rsch., Response to Request for Information, *supra* note 1.

[43] Advancing Racial Equity and Support for Underserved Communities Through the Federal Government, Exec. Order No. 13985, 86 Fed. Reg. 7009 (2021), www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/.

[44] KHOSHKHOO ET AL., *supra* note 1, at 51.

[45] *See* Tina J. Kauh, Jen'nan Ghazal Read & A. J. Scheitler, *The Critical Role of Racial/Ethnic Data Disaggregation for Health Equity* 40 POPULATION RSCH. & POL'Y REV. 2021, at 1, 2, 4–5, https://doi.org/10.1007/s11113-020-09631-6.

[46] Kwame Ture, Speech at University of California, Berkeley (Oct. 29, 1966) (transcript and audio available at *Stokely Carmichael: Speech at the University of California, Berkeley*, AM. RADIOWORKS, https://americanradioworks.publicradio.org/features/blackspeech/scarmichael.html); Brian Palmer, *When Did the Word Negro Become Taboo?*, SLATE (Jan. 11, 2010, 4:30 PM), https://slate.com/news-and-politics/2010/01/how-old-was-harry-reid-when-the-word-negro-became-taboo.html.

recommendation to remove this term from SPD 15. The Center also recommends ending the use of the term "majority" to refer to White people and "minority" to refer to ''Black or African American and Other Races,'' or ''All Other Races'' to denote collective descriptions of "minority" races. The use of "minority" to describe people that do not identify as White creates a distinction that establishes White as the norm and aggregates all other racialized populations as a monolithic aberration. Moreover, due to racial segregation in housing and education, the term "minority" does not reflect most people's experiences of living in the United States. Most people, regardless of how they identify racially, are the "majority" in the places where they live, work, and learn.[47] Moreover, less than 50 percent of the United States population is expected to identify as White by the next (2030) decennial census. As a result, the Center recommends that a revised SPD 15 discontinue use of the terms "majority" and "minority."

b. Remove the Terms "Far East" and "Asian Indian"

The "Far East" is an antiquated term that conveys imperial notions of exoticism. Terminology like "the East," "the Far East," and "the Orient" perpetuate racist stereotypes.[48] As a result, the Center recommends that future revisions to SPD 15 do not use the terms "the East," "the Far East," or "the Orient" in reference to East Asia. Additionally, while "Asian Indian" has been adopted in a few niche circumstances, it is not broadly used. The term aggregates populations with distinct cultural and socioeconomic differences (e.g., Pakistani, Nepali, and Bangladeshi)[49] and risks perpetuating a pattern of South Asian exclusion by confusing data between respondents who intend to highlight ties to India and those that intend to identify ties to the Indian subcontinent.[50] The Center recommends including more granular ethnic or national origin categories to mitigate improper aggregation.

---

[47] See generally THOMAS SUGRUE, ORIGINS OF THE URBAN CRISIS (2005); THE GEOGRAPHY OF OPPORTUNITY: RACE AND HOUSING CHOICE IN METROPOLITAN AMERICA (ed. Xavier Briggs 2005), www.brookings.edu/wp-content/uploads/2016/07/geographyofopportunity_chapter-1.pdf; MARTHA CECILIA BOTTIA, POVERTY & RACE RSCH. ACTION COUNCIL, RESIDENTIAL SEGREGATION AND IMMIGRANTS' OUTCOMES (2019), www.jstor.org/stable/resrep27199.10; Jeremy Pais, *Intergenerational Neighborhood Attainment and the Legacy of Racial Residential Segregation: A Causal Mediation Analysis*, 54 DEMOGRAPHY 1221 (2017), www.jstor.org/stable/45047294; Sequoia Carrillo & Pooja Salhotra, *The U.S. Student Population Is More Diverse, but Schools Are Still Highly Segregated*, NPR (July 14, 2022, 5:13 AM), www.npr.org/2022/07/14/1111060299/school-segregation-report.
[48] Christopher Hill, *What's the Matter with Saying 'The Orient'?*, JAPAN SOC'Y, https://aboutjapan.japansociety.org/content.cfm/whats_the_matter_with_saying_the_orient#sthash.Tzx 2MXtZ.7IsrONf5.dpbs (last visited Mar. 22, 2023); Sahar Aziz, Khaled A. Beydoun, Dalia Mogahed & Lakshmi Sridaran, *Islamophobia, in* BOS. UNIV. CTR. FOR ANTIRACIST RSCH., MOVING TOWARD ANTIBIGOTRY 145–46 (2022), www.bu.edu/antiracism-center/files/2022/06/Islamophobia.pdf.
[49] Vinay Harpalani, Desi*Crit: Theorizing the Racial Ambiguity of South Asian Americans*, 69 NYU Ann. Survey Am. L. 77 (2013), https://ssrn.com/abstract=2308892; Roopa Kalyanaraman Marcello, Johanna Dolle, Areeba Tariq, Sharanjit Kaur, Linda Wong, Joan Curcio, Rosy Thachil, Stella S. Yi & Nadia Islam, *Disaggregating Asian Race Reveals COVID-19 Disparities Among Asian American Patients at New York City's Public Hospital System*, 137 PUB. HEALTH REPS. 317 (2022), https://doi.org/10.1177/00333549211061313.
[50] Jennifer Lee & Karthick Ramakrishnan, *Who Counts as Asian*, 43 ETHNIC & RACIAL STUD. 1733-56 (2020), https://doi.org/10.1080/01419870.2019.1671600.

c. Maintain the Definition of AIAN

The Center recommends that the definition of AIAN not be expanded to include Indigenous populations from Latin America, so as not to obscure the particular experiences of the current AIAN population. Instead, the Center recommends including "Indigenous," among the ethnic and national origin categories, allowing respondents to identify as Indigenous regardless of which racial category they select. This would allow someone from Latin America to reflect that they are Indigenous irrespective of their selection of a racial category. The Center also recommends that OMB and the working group meaningfully engage with tribal leaders, AIAN communities, and other Indigenous communities in the United States to ensure accurate and comprehensive data collection and reporting.

d. Qualify the Use of "Multiracial," and Analyze Data Alone or In Combination

The Center recommends designating only respondents who select more than one racial category as "multiracial." As noted above, where people are designated as "multiracial" their data should be analyzed as a group but also as part of the respective racial groups with which respondents identify to avoid inaccuracies and dilution. For example, government agencies and researchers should both be able to aggregate "Black" alone, and "Black" in combination with any other racial category to properly evaluate data about people who identify as Black and better track anti-Black racism.

Respondents who select more than one ethnic or national origin category should be designated as selecting "more than one category" or some other term that avoids conflating ethnicity/national origin with the construct of race.

V. **OMB Should Conduct Additional Research to Improve Tracking and Addressing Racism**

The Center has several recommendations for further research that OMB can undertake— or encourage researchers to undertake—to continue to improve the standards used to collect data and better track racism. The following topics have serious implications for the administration of government programs and the ability to detect and address racism. The Center recommends that federal statistical agencies conduct targeted research to determine the best ways to address these issues and make further research-driven revisions to SPD 15 in the future.

a. Disaggregate the "Asian" Racial Category

Research shows that there are consistent racially disparate outcomes between the East Asian population compared to the South Asian and Southeast Asian populations.[51] Disaggregating the "Asian" racial category can reveal stark differences in socioeconomic status, health outcomes, and educational attainment between the East Asian population

---

[51] Neil G. Ruiz, Sunny Shao & Sono Shah, *What It Means to Be Asian in America*, Pew Rsch. Ctr. (Aug. 2, 2022), www.pewresearch.org/race-ethnicity/2022/08/02/what-it-means-to-be-asian-in-america/; Harpalani, *supra* note 49.

and the South Asian population.[52] This is especially important where the number of ethnicity or national origin categories listed on the question form are limited by administrability concerns.

Given the underdeveloped body of research investigating how disaggregating the "Asian" racial category could affect the rest of the data collection operation, the Center recommends comprehensive research into issues arising from using "East Asian," "South Asian," and "Southeast Asian" in the place of "Asian" for the minimum categories required under SPD 15.

b. Collect Data Using More Granular and Multidimensional Metrics

As data collection transitions to a digital medium, there is a growing opportunity to collect racial and ethnic data without the limitations of paper. Digital questionnaires offer the possibility of listing more comprehensive ethnic or national origin categories in a dropdown menu, allowing for more granular data collection.

There have also been calls for data collection reflecting multidimensional metrics of race, such as complexion, in order to more specifically assess racial disparities within racial categories.[53] Such data could also reveal manifestations of colorism, promoting a better understanding of how complexion impacts treatment in society and by the government. While there is some research on the relationship between complexion and outcomes such as income and education,[54] federal statistical agencies should further examine this issue in the context of collecting racial and ethnic identity data.

c. Protect Data Privacy

With respect to personal data, there is an emerging need for mechanisms protecting the rights to be forgotten and to revoke consent.[55] As the volume of data on each person increases, everyone becomes more vulnerable to the mishandling of personal data and privacy encroachment.[56] The Center recommends conducting research to identify best practices for implementing these protections. By exploring the potential benefits and challenges of incorporating the rights to be forgotten and to revoke consent, OMB can promote more equitable and responsible data collection.

d. Develop Tribal Data Sovereignty Agreements

There is an urgent need to develop a strategy at the federal level of government for preserving tribal data sovereignty and working collaboratively with Indigenous populations in data collection. OMB should develop formal agreements that address

---

[52] Margaret Simms, *"Model Minority" Myth Hides the Economic Realities of Many Asian Americans*, URBAN INST. (May 2, 2017), www.urban.org/urban-wire/model-minority-myth-hides-economic-realities-many-asian-americans; Kalyanaraman Marcello et al., *supra* note 49.

[53] López et al., *supra* note 33.

[54] Gonzalez et al., *supra* note 33.

[55] SEETA PEÑA GANGADHARAN, TAWANA PETTY, TAMIKA LEWIS & MARIELLA SABA, OUR DATA BODIES, DIGITAL DEFENSE PLAYBOOK: COMMUNITY POWER TOOLS FOR RECLAIMING DATA (2018), www.odbproject.org/wp-content/uploads/2019/03/ODB_DDP_HighRes_Single.pdf.

[56] *Id.*

ethical concerns and practical challenges to ensure that data collection respects the sovereignty of tribes and comprehensively illustrates the impact of policies on Indigenous communities. The development of such agreements should align with Indigenous values and contribute to more equitable outcomes.[57]

  e. Develop Equity-Based Metrics for Selecting the Ethnic/National Origin Categories Listed on the Question Form

We recommend further research regarding the ethnic and national origin categories presented as options on the questionnaire. In particular, research and testing should be conducted to (1) decouple ethnic categories from categories based on national origin, and (2) develop equity-based metrics to determine what categories to list, rather than selecting categories by population size.

Our proposed question (Figure A) mirrors the ethnicity and national origin categories recommended by the interagency working group (Figure 2), which were selected based on population size. This approach is an imperfect next step that is limited by a lack of researched and tested alternatives. We recommend further research into creating categories that distinguish between ethnicity and national origin. Additionally, we recommend research and development of equitable metrics for selecting the ethnicity and national origin categories that are listed on the question form, rather than relying on population size. For instance, the federal government has a responsibility to populations forced to migrate due to U.S. interventions abroad, such as people from Vietnam, Cambodia, Laos, Korea, Iraq, Afghanistan, and much of Central America. More robust research should account for these considerations and others that may counsel for tracking particular experiences of racism. As discussed below, these categories should also be regularly reviewed and amended.

  f. Implement Regular Review Process

The concept of race is continually constructed, and patterns in racial experience and racial identity are constantly changing.[58] Accordingly, OMB must institute a regular review process for SPD 15 to ensure that the categories closely approximate current conceptions of race, and thus enable the detection and mitigation of racism. The Center recommends a continuously improved and standardized review system for SPD 15 because improving the racial and ethnic categories used for data collection is not a one-time endeavor.

---

[57] Stephanie Russo Carroll, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, Kay Raseroka, Desi Rodriguez-Lonebear, Robyn Rowe, Rodrigo Sara, Jennifer D. Walker, Jane Anderson, & Maui Hudson, *The CARE Principles for Indigenous Data Governance*, 19 DATA SCI. J. 43 (2020), https://doi.org/10.5334/dsj-2020-043.

[58] Ann Morning, *Toward a Sociology of Racial Conceptualization for the 21st Century*, 87 SOC. FORCES 1167 (2009), https://doi.org/10.1353/sof.0.0169.