Data Collection Without Definitions

William A. Darity Jr.* Stephan Lefebvre†

December 12, 2024

Abstract

The Office of Management and Budget's Statistical Directive No. 15, first issued in 1977, revised in 1997, and revised again in 2024, sets minimum standards for federal government collection and reporting of data by race and ethnicity. We find that Directive 15 does accomplish its intended purpose of promoting data comparability and sharing across the government and beyond. Despite these benefits, the Directive is regularly the subject of criticism, particularly with regard to the definitions it provides for each of the seven racial and ethnic categories at its center. We describe problems with the definitions, including their internal flaws: they display circular logic, they are inconsistent, they conflate race with related concepts such as ancestry and nationality, and they inappropriately constrain the identity choices of individuals and groups. In this paper, we analyze a novel proposal: dispense with the definitions altogether. We find the proposed change to be narrowly tailored and pragmatic; it would immediately resolve problems with the definitions, increase data comparability across time, and increase flexibility provided under Directive 15.

Keywords: race data, demography, identity, federal statistics, Census

JEL Codes: Z13, C83, J15

*Duke University, Durham, NC, USA
†Bucknell University, Lewisburg, PA, USA

Contents

1 Introduction

Since at least 1977, the U.S. government has recognized the need for a coordinated, federal minimum standard for the collection and reporting of data by race and has promulgated these standards through the Office of Management and Budget's Statistical Directive No. 15 (OMB, 1978; OMB, 1997; OMB, 2024). Notwithstanding numerous benefits pursuant upon Directive 15, it includes a set of unworkable definitions for each of the racial and/or ethnic categories that unnecessarily constrain individual identification options, dilute a focus on race-based civil rights that was the Directive's impetus, and involve the federal government in a process of legitimation of racial hierarchy (Swartz, 1997). Building on previous critiques, we find that the Directive's definitions are inconsistent, circular, incomplete, and they inappropriately conflate race with concepts such as nationality and the geographical origin of hereditary ancestors.

In light of these problems, we analyze a novel proposal: dispense with the definitions altogether. We find that getting rid of the definitions for each of the racial and/or ethnic categories would have few negative consequences, it would overcome the problems of ill-defined racial groups, it would facilitate improved data collection and analysis by eliminating sources of confusion and inconsistency, and it would significantly lessen the federal government's role in ascribing characteristics to racial groups and racialized individuals. To vet the policy, we anticipate various counterarguments, finding, as a practical matter, the proposed change will not endanger data validity, nor comparability across time, nor data interpretability, but rather enhance data quality. In practice, the definitions are rarely included in surveys and they do not seem to be considered by most respondents to government questionnaires. Thus, there seems to be little to lose and much to gain. If implemented, this proposal would contribute to the federal government and the public's legitimate interests in monitoring outcomes by race.

Directive 15 specifies a minimum set of seven racial and ethnic categories and provides definitions that characterize membership for each group. The seven categories are: American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Middle Eastern or North African, Native Hawaiian or Pacific Islander, and White. Prior to the March 2024 update (see Table 1), there were five racial categories and the two ethnic categories, Hispanic or Latino and Not Hispanic or Latino.

Going forward, there will not be a distinction between race and ethnic categories; the new scheme will be one question for which individuals may indicate membership in multiple groups, with the possibility of checkboxes for detailed subcategories (e.g. Mexican, Puerto Rican, Cuban for Hispanic or Latino) and/or space for write-in responses.

Directive 15 provides additional guidance on working with federal race data. It specifies, for example, that racial groups be treated as "social-political constructs" (OMB, 2024, p. 22191), it recommends self-identification of race whenever possible, and provides an example of specific wording for the race and ethnicity question. More detailed information can be collected so long as the data can be aggregated into the main seven categories listed in the Directive. These standards are designed to be used in all federally sponsored data collection or reporting that includes race or ethnicity, general federal program administrative and grant reporting, and civil rights and other compliance reporting.

The definitions have a number of problems. Table 1 reproduces the full text of the original 1977 definitions alongside the revisions made in 1997 and 2024. For an example of inconsistency across the categories and circularity within the definitions, consider the category Black or African American. The definition reads, in part, "Individuals with origins in any of the Black racial groups of Africa," making it the only category that specifies a color in the definition.

For an example of lack of comprehensiveness, which we interpret to mean that some potential respondents do not have at least one category that is appropriate for them, consider the "American Indian or Alaska Native" category which, in the 1977 version, did not include indigenous people from South America. The definition was changed in 1997 to include descendants of "the original peoples of... South America," but the condition that individuals "maintain[ ] tribal affiliation or community attachment" was added.

This reflects specific legal and citizenship implications of tribal affiliation in the U.S., and indicates problems with comprehensiveness were only partially addressed since it is not clear that such criteria should apply to all indigenous groups in the Americas. The latest revision dispenses with the "tribal affiliation or community attachment" requirement and instead lists examples of groups, some of which are federally recognized tribes in the U.S. This is significant because each tribal nation has latitude in determining criterion for membership that may or may not align with the Directive 15 definition.

Our critique of the definitions is not based on the idea that racial group membership is so amorphous or subjective that it inherently presents problems for defining terms or quantifying dimensions of racial hierarchy. Rather, we argue that the current definitions are unworkable and that any similar attempt to define membership in broad racial groups for uniform use as a minimum standard by the federal government and by industries with significant ties to the federal government (healthcare, education, law enforcement, etc.) is doomed to fail.

The social operation of race cannot be reduced to definitions for the expansive purposes of Directive 15. The proposal to remove the definitions is limited in scope and does not constitute a comprehensive plan for revising the Directive. It is possible to support the proposed change as a move towards better data collection regardless of one's position on a number of other issues related to Directive 15. Removing the definitions is not a sign that we should throw up our hands and conclude quantitative data on race is uniquely difficult.

The general problem of collecting information about an individuals' race is complex. Race is a multidimensional concept (Roth, 2016). Some of these dimensions are phenotype, self-identity, social identity, legal standards (which ultimately are "community standards") and family history. Any of these dimensions may be geographically and temporally contingent, and they may have extensive and intensive margins.

The specific problem of designing and implementing federal minimum standards is of even greater complexity. "Federal" means that the standards need to be designed with a full appreciation of both the country's diversity and the diverse needs of hundreds of federal agencies. The standards should be designed considering every conceivable person who may find themselves filling out a federal government form or survey. "Minimum" means the standards need to be designed with the full complexity of race in mind, while pairing down the guidance to its most succinct and economical form, leaving as much flexibility as possible for different use cases.

The main contribution of this paper is to analyze a novel policy proposal. The latest revision of Directive 15 involved a sprawling interagency process that leveraged the Census' extensive 2015 National Content Test (Mathews et al., 2017) and involved "[collecting] over 20,000 comments. . . , 94 separate 30-minute listening sessions, and" virtual town halls where "3,350 people joined [and] where over 200 people spoke" (OMB, 2024, p. 22184). This paper's scope is limited to the definitions included in Directive 15 and considers a proposal, to the best of our knowledge, that has not been

considered in the studies above or elsewhere.

Our paper is also distinct in that it combines theories of racialization with research on the practical realities of policy change around Directive 15 specifically and survey methodology more broadly. Other researchers have critiqued the categories and definitions of Directive 15 (Tamayo Lott, 1997), particularly in the realms of law (Toro, 1995) and health (Wallman, Evinger, & Schechter, 2000). However, these works do not analyze alternatives. There is little academic work analyzing practical policy changes to address the problems we echo in this paper. For example, Hernández (1998, p. 167) contains a critique of the definitions and an interesting example of how to elicit individual race identity in a survey (a lengthy prompt to "invite[ ] responses about personal identity"), but it is presented as "a vehicle for initiating... discussion... rather than being a concrete model for statistical data collection."

Gimenez (1989) argues that "any standardized terminology" for racial groups "is unavoidably flawed and conducive to the development of racist... analysis of the data thus produced," but her proposal for a new question only discusses Latinxs. Gimenez's proposed solution to disaggregate Latinxs into six[1] categories is not qualitatively dissimilar to the current standards and echos other calls for data disaggregation or, similarly, abandoning the pan-ethnic category "Hispanic or Latino."

The proposal we analyze is consistent with many different levels of (dis)aggregation, and we express no preference or analysis for how many groups or which groups should be present in Directive 15. The proposal's limited scope reflects our interest in practicality for implementation. As Skerry (2000, p. 43) reminds us, this does not imply that the seemingly small change lacks importance because "The most contentious boundaries that the Census Bureau has to deal with are those concerning race and ethnicity."

In Section 2, we present an overview of Directive 15's contents, its policy history, as well as its reach and lack of enforcement mechanism. In Section 3, we argue that the definitions for the racial groups used in the Directive are inconsistent, circular, incomplete, unhelpful because they conflate race with other demographic concepts, and inappropriate, insofar as the federal government is, however unintentionally, authoritatively assigning particular racial groups with characteristics that are not accepted by all members of that group. Section 4 contains details of the proposal and

[1]The six groups proposed in Gimenez (1989) are Mexicans, Puerto Ricans, Cubans, Central American refugees, Central American immigrants, and South American immigrants.

Section 5 provides an analysis of its benefits, potential drawbacks, and implications.

2 Contents of OMB Directive 15

Directive 15 (OMB, 2024) defines seven race and ethnic categories and prescribes these as minimum categories for data collection by the federal government. The categories are: American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Middle Eastern or North African, Native Hawaiian or Pacific Islander, and White. As of the 1997 version, the standards require that respondents be able to select more than one racial category. In addition to setting the number and names of the minimum categories, Directive 15 defines each of the categories. Table 1 reproduces definitions from the original 1977 standards, the first revision made in 1997, and the second and latest Directive 15 text issued in March 2024.

Successive changes to Directive 15, namely increasing the number of categories, highlight the contradictions of using a scheme that was intended for use in race-based civil rights but which currently erases distinctions between groups that are racialized differently in the U.S. and abroad (e.g. Palestinians and Israelis, Indians and Chinese, Turks and Kurds). The standards do not preclude collecting data with finer gradations, but any such collection must be done in such a way as to allow the aggregation of individuals into one or more of the seven official categories.

Directive 15 states its purpose for "Federal statistics, program administrative reporting, and civil rights compliance reporting" (OMB, 2024, p. 22191), but does not specify the circumstances under which the definitions are to be used. For example, it is not clear whether the definitions should be provided to survey respondents; if provided, it is not clear whether they should be visible alongside the category options, included as a footnote, linked to as an online reference, or something else. To our knowledge, no major national surveys provide the definitions from Directive 15 alongside the questions on race and ethnicity.

On the rare occasions that the definitions do appear outside of direct commentary on Directive 15, they are reproduced as supplemental background information and provided alongside reported statistics after the fact, not in connection with the survey instrument. In some cases, the data is aggregated from surveys that ask more detailed questions on race and so the definitions were likely not used in data collection, raising questions about the appropriateness of applying the definitions.

Directive 15 provides additional guidance around the nature of race, different question formats, and which body is authorized to provide exceptions to the standards. Since 1977, the standards contain similar language that: (1) the categories are "sociopolitical constructs" that "are not an attempt to define race and ethnicity biologically or genetically" (OMB, 2024, p. 22183), (2) the categories "are not to be used as determinants of eligibility for participation in any Federal program" (OMB, 2024, p. 22183), and (3) any deviation from the standards requires specific authorization by the Office of Management and Budget. The latest standards establish that (4) race and ethnicity should be collected in a one question format and that (5) "the seven minimum race and ethnicity categories shall be treated co-equally.... collection forms may not indicate to respondents that they should interpret some categories as ethnicities and others as races, or otherwise indicate conceptual differences among the minimum categories" (OMB, 2024, p. 22195). The last item presents a change from previous Directive 15 standards that emphasized a conceptual difference between Hispanics or Latinos, who could be of any race and were defined by reference to "Spanish culture or origin."

2.1 Policy history

Statistical Directive 15, first issued in 1977, arose in response to a growing need for data on race and ethnicity for civil rights monitoring and enforcement (Federal Interagency Committee on Education, 1975). Agencies within the federal government collect data as part of their mission or due to legal mandates. For example, both the Civil Rights Act of 1964 and the Voting Rights Act of 1965 require specific agencies to collect data on race and ethnicity. The 1977 standards created a new scheme for collecting racial data that required, at minimum, four categories or the ability to aggregate more detailed data into the four broad categories, which were American Indian or Alaska Native, Asian or Pacific Islander, Black, and White.

The 1977 Directive 15 implemented the previous year's Public Law 94-311, which required the federal government to collect information to make Hispanics as a group identifiable in the nation's statistical reports (94th Congress, 1976). This mandate led to a two-part question, whereby respondents are asked to identify as Hispanic or non-Hispanic in one question and asked to identify their race in a second question. Public Law 94-311 was proposed and passed in Congress largely due to the efforts of Representative Edward Roybal of California and the newly formed Congressional Hispanic Caucus (Robbin, 2000).

The first interagency process for updating and revising Directive 15 began in 1993 (Citro, 1997). Official reports and statements from government officials frequently cite the need to "reflect[] the increasing diversity of our Nation's population that has resulted primarily from growth in immigration and interracial marriages" (Wallman, 1998, p. 31). For example, no experts in American Indian affairs, community or political representatives, were involved in drafting the original Directive 15 (Forbes, 1992). The 1997 update to Directive 15 (1) split the "Asian or Pacific Islander" group into "Asian" and "Native Hawaiian or Other Pacific Islander," thus establishing five racial categories as opposed to four, (2) allowed individuals to select (or be assigned to) more than one racial group, (3) edited the definitions to include indigenous peoples from South and Central America in the Native American category, and (4) changed the category names and definitions for some of the groups, altering the list of country examples in the Asian group and adding the term Latino to the Hispanic category. As with Public Law 94-311, the separation of Native Hawaiians from Asians and from Native Americans was achieved in part through the advocacy of a particular congressman, Senator Daniel Akaka of Hawaii. The original 1977 and the revised 1997 version of Directive 15 both represent the outcome of intergroup and intragroup cooperation and competition, a complex interagency process, and numerous compromises with uneven representation of expertise.

The U.S. Census Bureau is subject to Directive 15. Therefore, the millennial Census operated under the then three-year-old, revised Directive 15 with one high profile exception that allows us to illustrate how the standards work in practice. The standards allow for significant flexibility, including the selection of more than one category and including any number of detailed categories for race and ethnicity, so long as it is possible to aggregate up to the broad Directive 15 categories.

The Census Bureau requested and was granted an exemption to this requirement in order to add a sixth category, "Some other race."[2] This category is problematic under Directive 15 because those indicating "Some other race" with no other racial category could not be aggregated into one or even multiple of the five official racial categories. In the 2000 Census, 97 percent of those who indicated Some other race also identified as Hispanic or Latino. As Census officials indicate, the Bureau has as an explicit goal "to reduce reporting of 'Other race' by Hispanics in the race item" (Martin, Demaio, & Campanelli, 1990, p. 551). Although beyond the scope of this paper, we note that the Census Bureau has spent the past 35 years struggling to collect race data on Hispanics.

[2]The residual category "Other" has been in use by the Census since 1910.

Since 1997, the most contentious issues involve (1) allowing an explicit "Multiracial" category to "directly promote a distinct multiracial identity" (Hernández, 1998, p. 100), (2) combining the race and ethnicity question, which reverses decades long messaging that Hispanics are not a race but a pan-ethnic category incorporating people of any race,[3] (3) accurately reflecting the identities of American Indian and Native Hawaiians, and (4) creating a Middle East and North Africa category that includes individuals who, in the past, would have fallen under the definitional category of whites (see Table 1).

The latest revision of Directive 15 is only the second in nearly 50 years. It was led by the Federal Interagency Technical Working Group on Race and Ethnicity Standards, a combination of the 13 Principle Statistical Agencies along with 25 additional agencies selected for their reliance on race and ethnicity data (Office of Management and Budget, 2023). This 2024 revision added a new category, Middle Eastern or North African, edited each of the category definitions, removed the word "Other" from the category name Native Hawaiian or Other Pacific Islander, and contained several other changes in guidance discussed later in this paper.

The revision processes have not been construed narrowly, signaling a need to refocus on the purpose and effects of Directive 15. In particular, the process has increasingly focused on reflecting the nation's "diversity," though diversity of what is not well defined. In public comment and instrument testing, the focus on ethnic or cultural identities has not explicitly tied to the civil rights monitoring or other of Directive 15's original purposes discussed below. These observations are made not to suggest that these topics are not important in many contexts, but to suggest that they may be of only minor relevance for a minimum standard for all data collection undertaken by the federal government.

Data are generally more comparable over time if changes to statistical standards are infrequent and modest. Standards have changed only twice since their adoption, despite generally recognized problems with the definitions in part due to the costliness and contentiousness of the revision process.

2.2 Stated purpose, reach, and enforcement

In this section, we review the purpose, reach, and enforcement mechanisms of Directive 15. Both

[3] As a counterpoint, respondents are still able to select a traditional race category along with Hispanic or Latino.

the original and the revised Directive 15 contain language about its intended use: the minimum categories are to be used "by all agencies... for civil rights compliance reporting and equal employment reporting for both the public and private sectors and for all levels of government" and for all "federally sponsored statistical data collection where race and/or ethnicity is required" (OMB, 1978).

The federal minimum standards around race and ethnicity set forth in Directive 15 provide a number of benefits. They help to encourage similar practices around data across the federal government and related bodies, which facilitates sharing and comparing data across official statistical reports. The standards contribute to government functioning through increased efficiency: while each department and agency is charged with planning and executing its own strategies around data, the guidance, particularly around this contentious and contested issue, reduces duplication of efforts.

The standards do not preclude data collection that goes beyond Directive 15's broad racial and ethnic categories nor does it limit approaches to studying racial and ethnic inequality to particular multiple choice questions or survey methods. Agencies are free to experiment or to pursue research that is based on qualitative approaches or alternative quantitative methodologies. Directive 15 is intended to provide *minimum standards* for regular agency functions.

Whether intended or not, comprehensive use of these standards by the federal government has a trickle-down effect. Directive 15 influences the operations of organizations and institutions with significant contact with the federal government. OMB recognizes this in official communications: "Even where it is not required, Directive No. 15 standards are often used in State and business record systems and by markets as a matter of convenience and to facilitate comparisons with other data sets" (OMB, 1995, p. 44676).

Insofar as Directive 15 shapes administrative practices throughout the country, nearly everyone in the country fills out a form or survey where they are asked to fit their identity into its categories. The far-reaching nature of Directive 15 is a key consideration for any proposed changes: "Any changes in Directive No. 15 will be imposed on tens of thousands of State and local agencies such as law enforcement agencies (through the Uniform Crime Reporting system), school districts, the business community, and others required to use the Directive in reporting these data to the Federal government" (OMB, 1995, p. 44680).

In Section 5, we will argue that removing the definitions would not have negative consequences even when taking an expansive view of its reach and influence. While Directive 15 has a number of important benefits, none of these depend on the definitions. Part of the reason that the proposed change would not be disruptive has to do with which aspects of Directive 15 are the most influential on practices within the agencies, institutions, and organizations listed above.

The influence of Directive 15 is most directly seen in which racial and ethnic categories are available in surveys. The current standards require that individuals can be aggregated into seven specific race and/or ethnic categories. Surveys in compliance with Directive 15 provide these categories or they provide additional options that are clearly subcategories. There is little evidence that the definitions provided in Directive 15 are used in practice as they are rarely cited or reproduced and never included in surveys for respondents to refer to as they make their selections.

The first update to Directive 15 required setting forth guiding principles for the revision process. There were 12 principles, including "Respect for individual dignity" (OMB, 1994). Interestingly, the Principles, originally published in 1994, were reissued in 1995 with small changes that reveal some of the thinking or intended purpose of Directive 15. For example, the later version reproduces Principle 5, "Foremost consideration should be given to data aggregations by race and ethnicity that are useful for statistical analysis, program administration and assessment, and enforcement of existing laws and judicial decisions," but removes reference to, "enforcement of existing laws and judicial decisions."

Similar changes were made to limit the scope of Principle 6, "While Federal data needs for racial and ethnic data are of primary importance, consideration should also be given to needs at the State and local government levels, including American Indian tribal and Alaska Native village governments," which was edited to subordinate more clearly non-federal data needs.

Directive 15 is far reaching, both by explicit intention and in practical effect. Any critique of proposed changes to Directive 15 should account for the fact that the purpose of Directive 15 is not to define racial groups for individuals nor to provide guidance for researchers doing work on race apart from the design of federal government surveys, administrative forms, and reports.

Moreover, there is no clear enforcement mechanism for Directive 15. In our exploration we found no evidence of active enforcement efforts. The Directive does not specify when race and ethnicity data should be collected. It specifies the category names and definitions but does not

provide any explanation regarding how the definitions are to be used.

The effect of Directive 15 is to promote a lower limit of data comparability and collaboration between agencies. This does not guarantee data on race and ethnicity are easily comparable across federal statistical agencies. One study, though not at the federal level, illustrates the current landscape. It found that data on race was not comparable across states or even within a state across different survey modalities. The State Health Access Data Assistance Center at the University of Minnesota (2021) found large discrepancies from state to state and within states between online and paper versions of enrollment forms for Medicaid applications. Of the states in their sample, fourteen offered more racial category options online, nine offered more on the paper version, and ten offered the same number on both forms.

Policy guidance outside of Directive 15 significantly affects how the standards function in practice. For example, an OMB Bulletin from March 2000 states that "Responses that combine one minority race and white are allocated to the minority race" (Lew, 2000). That is, statistical practices arguably consistent with Directive 15 can follow rules of hypodescent in which whiteness is treated as "pure" so that children with both white and non-white parents are assigned their non-white parent's race. This is significant in several ways (Myers Jr., 1997), but for our argument the key point is this: the definitions provided in the text of Directive 15 are not the only nor even the most important way in which the federal government provides guidance about how to work with data on race. In other instances, the Census recoded respondents who indicated Hispanic into the "Some other race" group (Schuster, 2021) and recoded individuals to native American when they include particular write-in tribal categories (e.g., Cherokee), even when the survey respondent did not check the box for native American (Eschbach & Taylor, in press).

In other words, Directive 15 is not the superseding or sole source of information for understanding what race categories mean; the categories are defined implicitly by recommended and implemented practices such as the rule of hypodescent, though these may contradict the definitions. This means that removing the definitions has the potential to end the problems associated with them. We know of no alternative definitions or way of writing alternative definitions that meet the necessary criteria for Directive 15.

The problem at issue in this paper does not involve questions about how the U.S. government should ideally or usually collect data on race. Those questions do not recognize that Directive

15 only specifies *minimuim* standards. As is acknowledged in the text setting out the revision principles, "The agencies recognize that these principles may in some cases represent competing goals for the standard.... Through the review process, it will be necessary to balance statistical issues, needs for data, and social concerns" (OMB, 1994, p. 29834). In Section 5 we discuss our findings that the proposed change would bring Directive 15 closer to fulfilling multiple of the revision principles currently in discussion, with little to no down side in terms of the other principles.

3 Problems with the federal government's definitions

In this section, we identify several major problems with the definitions for racial groups as they are written in the latest Directive 15 text. The problems we raise are not new to the literature; many have been advanced as critiques of racial classification schemes without reference to Directive 15, others have been levied against Directive 15 specifically, and many have discussed in official government reports as part of various revision processes. Our argument is unique because we only deal with the definitions for racial groups as part of a set of minimum standards for statistical practices and because we evaluate a practical solution that we have determined would immediately resolve all of the problems discussed in this section, could feasibly be implemented, and is compatible with many other modifications to the Directive.

As part of the 1997 revision process, the U.S. government acknowledged that "the categories in Directive No. 15 confuse some respondents because they are inconsistent, too broad for some purposes, and the concepts of race, Hispanic origin, and ancestry overlap" (OMB, 1995, p. 44677). Despite this admission, Table 1 shows that only modest changes were made in each revision and that many of the problems remain.

The original categories from the 1977 Directive 15, largely unchanged in the 1997 revision, were presented using a great deal of research and administrative jargon, but they clearly conform to anthropologically discredited ideas of "the ethno-racial pentagon" (Hollinger, 2000). Under different names, the five categories would be familiar to educated Americans of the 19th century: white (European), black (African), red (Native American), brown (South American), and yellow (East Asian). It seems to be impossible to craft definitions for these specific categories that (1) fulfill the stated purposes of Directive 15, (2) provide sufficient uniformity for their countless federal

applications, (3) capture the real-world asymmetries in how racial groups function, and (4) respect individual and group identity choice. Here, we discuss problems with the current definitions.

3.1 Inconsistent, circular, and noncomprehensive

The definitions provided in Directive 15 are inconsistent, noncomprehensive, and, in some cases, weak due to circularity. Among the many inconsistencies, consider the following three examples.

First, if black is a skin color, it is only referenced for the Black or African American group. This group is also the only racial category that does not contain the phrase "any of the original peoples of" in reference to a geographical origin which, given the diversity and intensity of centuries-old colonial projects in Africa, functions to deemphasize African claims to particular land and resources.

Second, the category American Indian or Alaska Native was the only category requiring specific cultural and political practices ("maintains tribal affiliation or community attachment"), though this qualifier was officially removed in the latest revision. The added burden to identification included for American Indian or Alaska Natives in Directive 15 functionally served to reduce the size of this group, consistent with the project of settler colonialism.[4]

There is not yet enough evidence yet to reach conclusions about the March 2024 formulation. Again, we do not argue that the qualification "maintains tribal affiliation or community attachment" is never appropriate or that removing that qualification by itself would improve the definitions. Our argument is that the definitions are flawed and unnecessary for achieving Directive 15's intended purpose. The last and certainly the most humorous of the inconsistencies was the typographical error in the 1997 version wherein the term "Cubans" appears twice in the definition of Hispanic or Latino, corrected in the latest revision.

Consider the definition of the category Black or African American, "Individuals with origins in any of the Black racial groups of Africa." No independent definition of what it means to be black is offered. Anyone who has origins in a non-black (also undefined since "black" itself is undefined) racial group of Africa is presumed to have origins in some other continent or place. Indeed, the definitions are silent about any racial groups in Africa other than "Black racial groups."

Unlike most of the other racial groups in the classification scheme, the inclusion of the phrase

[4]Forbes (1992, p. 63) writes, "one reason for the directive's treatment of Native Americans was to reduce the numbers of persons qualifying for federal services as Indians."

"any of the Black racial groups" shifts this racial category from being exclusively based upon geographic origin (see Table 1). Oddly enough, the definition of white does not include the qualifier that individuals with European ancestry must have origins in any of the white racial groups, and this was true even when the white category included those from the Middle East and North Africa.

Even if parallelism were achieved—even if the wording of each definition were edited to be more consistent—that would not mean the categories would be more accurate or the meanings would be similar given the different historical contexts relevant for each group. Some self-identified blacks who have been in the United States for generations have no record of where in Africa their ancestors were born and do not wish to be called "African Americans."

Historically, African Americans as a group have been subject to campaigns intended to suppress African culture and Afrocentric identities and to devalue links with Africa. In short, there are reasons why many blacks in the U.S. might or might not want to emphasize a connection with Africa. Removing the problematic definition would not preclude any self- and group-identification choices but would beneficially eliminate the symbolic imposition of a unified definition through government statistics.

Though we argue that the definitions are inconsistent, we recognize that hierarchy and unequal power is inherent in characterizing racial groups in all societies in which they have been studied. The way race functions since European colonization of North America has been to create hierarchies and, correspondingly, conventional definitions of who is white, black, non-white or non-black do not function symmetrically.

This operates at every level, including the social rules that characterize membership in racial groups. As Fields (1982, p. 149) writes, there is "a well-known anomaly of American racial convention that considers a white woman capable of giving birth to a black child but denies that a black woman can give birth to a white child."

Not only are the definitions inconsistent, but they are inconsistent without reason or rationale unsupportive of nineteenth century constructs of race. Reference to the "Black racial groups of Africa" with no similar reference to the "White racial groups of Europe" in the white category definition does not conform, to the best of our knowledge, to any substantive current socio-political and scientific understanding of race (for current conceptualizations of race, see American Anthropological Association, 1998; American Association of Biological Anthropologists, 1996).

The process of collecting survey information can create a false appearance of similarity or parallelism. Any individual may have one or more racial identities in the United States, but race is not only nor even primarily an individual characteristic. Rather, it is a sociopolitical, economic, and ideological phenomenon that has evolved over time with geospatial variation.

Racial groups have specific histories, such as the relationship between whites and blacks during and in the wake of chattel capitalism or the relationship between whites and American Indians during and in the wake of the genocidal completion of the nation's settler colonial project. The historical and hierarchical characteristics of race as a social group phenomenon mean that creating a comprehensive list of definitions for individual membership is besides the point; definitions are not necessary to ensure meaningful data collection or any of the other goals of Directive 15.

Another inconsistency with respect to the Directive 15 definitions involves the legal meanings and implications of race for U.S. jurisprudence. For the first 89 years of the nation's existence an individual's race determined whether one was considered property or capable of owning other humans. Today, an individual's race is central for anti-discrimination case law (Desautels-Stein, 2012).

The court system typically considers a number of criteria when seeking to understand an individual's race, including "physical characteristics, documentary evidence, community perception, and expressive self-identification" (Yang, 2005, p. 406). In anti-discrimination cases, the perception of an individual is a major consideration and the "truth" of an individual's heritage or other ancestral characteristics is largely irrelevant.

The federal government defines membership in racial groups narrowly and specifically based on whether an individual has "origins in any of the original peoples" of various geographically defined regions (with notable exceptions that we have discussed), but for the purposes of applying the nation's laws, a variety of other standards are used (Ford, 1994).

The distinction is not just that the law is more nuanced, which would be an expected difference between the use of a concept in civil or criminal case law as opposed to questions on a survey or administrative form. Instead, the concepts contained in the Directive 15 definitions carry virtually no weight legally and, conversely, little of the sociopolitical definition of race is captured by geographically based category definitions.

Interestingly, there is a new, recursive type of circularity in the definitions that was not present

for the original Directive 15 that arises because the standards are nearly five decades old. The goal of the category names and definitions is to reflect social practice, to give people options that are close to how they might identify if given an open-ended response option, but social practice itself has been shaped by Directive 15.

Revisions to Directive 15 must account for changes in the social meaning of racial groups while preserving data comparability over time. However, *those revisions also have shaped how people think of themselves*. The categories and their associated meanings, if not the precise definitions, take on a legitimacy and authority of their own. They influence social change through a feedback loop.

Consider the Hispanic ethnicity category, introduced into federal data collection in Census 1970. For many today, it would be sensible to refer to Brazilians as Latinos/Latinas/Latinxs (Margolis, 1995). The particular recognition given to Spanish speakers that arose from a specific political climate reified them as an ethnic group through the administrative practices of the U.S. government (Fears, 2003). The 1977 Directive 15 implemented the previous year's Public Law 94-311, which required the federal government to collect information to make Hispanics as a group identifiable in the nation's statistical reports (94th Congress, 1976). This mandate led to a two-part question whereby respondents are asked to identify as Hispanic or non-Hispanic in one question and asked to identify their race in a second question. Public Law 94-311 was proposed and passed in Congress largely due to the efforts of Representative Edward Roybal of California and the newly formed Congressional Hispanic Caucus (Robbin, 2000).

Today, defining Latinidad in the U.S., a former colony and current colonial power, through reference to another colonial power, Spain, is not something that all Latinxs embrace (Rosa & Flores, 2017).

Public Law 94-311 was part of a flurry of activity following the 1964 Civil Rights Act. Seeing the legislative successes of the black community, various Latino groups interested in consolidating power realized that by joining the Mexican American community of the west and southwest, the Puerto Rican community in New York, and the Cuban American community in South Florida, their numbers could be compared to that of African Americans.

The outcome of these efforts is a reified understanding of Latinos in the U.S. conceptually distinct from blackness despite the groups' interconnected histories of racism and slavery throughout

the Americas. Therefore, reciprocally, revisions to the categories in Directive 15 are influenced by social phenomena that are themselves affected by Directive 15's definitions.

Furthermore, categories are not complete. Where does one situate individuals racially whose ancestors are Australian aborigines or New Zealand Maoris or any of the more than 40 indigenous Arctic zone communities (Walter & Anderson, 2013)? Given that much of the list is organized by countries, what specific countries constitute North Africa for the purposes of distinguishing the Middle Eastern or North African from the Black or African American category? Is the concept of "Black racial groups of Africa" an accepted construct from the perspective of immigrants from the African continent whose families arrived afted the 1960s when they find themselves in the U.S. filling out a form?

The definitions are based on the false premise that a particular set of racial meanings, however familiar they may be in the U.S., is appropriate for application through the federal government's minimum standards for race and ethnicity statistics. Although the revision process involves dozens of agencies, costly survey instrument testing, focus groups, and contentious consultations, anyone with access to a map and an internet search browser can find problems with the exhaustiveness of the definitions.

The categories are also incomplete when it comes to accounting for racialization as Chicano/ Chicana/Chicanx (Martinez, 1997; Toro, 1995). Sociological evidence shows that many Latinxs think of themselves as *racially* Latino (Hitlin, Brown, & Elder, 2007), and for good reason.

There are histories of racial epithets, discriminatory policies and practices, lynchings and other extrajudicial killings directed against American s with Mexican and Central American heritage (Mirandé, 2019; Shadowen, 2018). Latinxs are frequently racialized as "slightly tan, with dark hair and eyes" in media (see Rodríguez, 1997, for a book-length elaboration of the visual language around Latinidad), and there is a two-way racialization of "looking Latinx" and "sounding Latinx" involving speaking Spanish or Spanish-accented English (Rosa, 2019).

Yet, not all Latinxs are racialized as "brown." Collapsing Latinidad into a single-race group privileges phenotypically white Latinxs. A specific proposal for collecting race information on Latinidad must confront the highly successful campaign, of which the 1977 Directive 15 was a major part, to create a new social panethnicity called Hispanic (Mora, 2014). Such a proposal is beyond the scope of the current paper, but the significance of this issue for this type of proposal empha-

sizes problems with the current definitions and recognizes that various ways of better collecting information on Latinxs' racialization would be consistent with removing definitions from Directive 15.

For example, it would be possible to list race as a sociopolitical concept on surveys and within the Directive, while asking questions about identity or what López and Hogan (2021, p. 1) term "street-race," the typical category others assign you to "if you were walking down the street." Their proposed survey question includes the category "brown" as an option and, crucially, *does not provide definitions for the individual categories.*

The OMB has stated, "In line with the subjective nature of the concept, research shows people change how they classify themselves with respect to race and ethnicity" over time (OMB, 1995, p. 44677). This is not entirely accurate. Race is a multidimensional concept and some dimensions are more "subjective" than others; certainly, group-based phenomena such as the realities of race-based slavery and its aftermath are not completely "subjective."

The critiques we discuss in this section demonstrate how difficult it will be to fix the definitions in Directive 15 through edits. Given its wide-ranging intended purpose, no set of definitions we can conceive would not be subject to the critiques we raise. Still, our critique must not be taken as a proposal to eliminate federal race categories altogether. We have written extensively in favor of race-conscious policies, including reparations for U.S. descendants of slavery, and we have detailed the identity and ancestry requirements that could accompany such a policy (Darity & Mullen, 2020). We are advocating maintenance of the use of categories modified in accordance with changing social standards coupled with elimination of official definitions of the categories.

3.2 Conflating race with related concepts

In this section, we review problems with the definitions stemming from how Directive 15 unhelpfully conflates race with related concepts such as geographical origin by region (usually continents), nationality, ancestry, tribal affiliation, and ethnic identity. Reading the definitions, one is struck by a repeated use of the phrase, "Individuals with origins in any of the original peoples of. . . ," which appears in five of the seven category definitions.

An implicit power law is operative when accounting for generations in human lineages. Assuming each individual has two parents and counting back by generations we obtain four grandparents,

eight great-grandparents, then 16, 32, 64, and 128 in still earlier generations. Going back to the year 1619, approximately 13 generations ago, we would be tracking the "origins" (not the location) of 8,192 individuals.

Taken literally, this is patently unworkable. It is also not how people think about their racial identity. What does it mean for a person to "have origins" in a "original peoples" of a broad, ill-defined, contested geographical area? There are no sensible answers to these questions. We find that the definitions should be abandoned as inaccurate, misleading, and imprecise.

Although the original framing of Directive 15 and the 1997 revision insist that race is a sociopolitical concept, the definitions do not reflect this. The definitions in Directive 15 are ahistorical in two ways. First, they do not allow for changes in the meaning of racial categories over time. Instead they refer to fixed geographical referents consistent with 19th century hierarchies of "race." Second, they do not conform to the historical and contemporary ways in which race functions socially. The definitions support the myth of a biological (consanguine) construction of race. They draw on historically and anthropologically inaccurate assumptions about different groups as having discrete origins in fixed geographical areas.

The definitions and discussions around the revision process work to naturalize the idea of racial groups originating in discrete continents (Africa, Europe, North and South America) and seem to take for granted that individuals with a presumed ancestry from the same country share the same race. Or, incorporating logic implied by the phrase "original peoples," if we go back far enough all the peoples in a given country's territory would be identified in the same "race" today. This is a fiction. It is zombie anthropology: disavowed by physical anthropologists today but difficult to extricate from our cultural and governmental administrative practices (Fuentes et al., 2019).

There are many examples that illustrate the error in this framework. Filipinos are categorized as Asians under Directive 15, but the earliest known evidence of human inhabitants in the Philippines were people from Africa. These are the ancestors of various (typically) dark-skinned, contemporary Filipino ethnic groups (Larena et al., 2021).[5] To state the obvious, the original peoples of Asia and elsewhere likely came from Africa because *homo sapiens* first evolved in what we call present-day Africa. Skin did not lighten immediately upon crossing into Europe 50,000 years ago. Evidence

[5]This is not to suggest that some Filipino ethnic groups are not Asian, but that there are problems with the fundamental concepts used in the Directive 15 definitions.

indicates that lighter skin may have emerged later than previously thought, perhaps only eight thousand years ago (Mathieson et al., 2015).

There is a link between our critique of the definitions, which point to consanguinity (ancestry) as an important determinant of race, and critiques of race as a biological concept. In a biomedical research context, A. M. Johnson Jr. (2008) writes: "It is not their race that predisposes [African Americans] to hypertension, but their discriminatory treatment that is predicated on their grosser morphological traits. I cannot believe that there is anything in their genetic makeup that predisposes 'blacks' to have a higher rate of hypertension, because there is no identifiable genetic definition of a black person, only a societally constructed, morphologically based definition." Indeed, race is not primarily an individual characteristic. We support calls for more analysis of racialized organizations (Ray, 2019), institutions, systems, and phenomena based on data that are not limited to individual reports on race.

The problem of mixing concepts of race with geographical regions is apparent in the contentious issue, yet only partially resolved, of adding a separate Middle Eastern and North African (MENA) designation. Prior to March 2024, individuals "having origins in any of the original peoples of. . . the Middle East[ ] or North Africa" were classified as white. There is now a MENA category, notwithstanding disagreement among individuals who might be located in this group about this term. A common refrain asks, for example, middle of or east of what?[6] There is no agreement about the boundaries for this region.

Lack of a consistent understanding is found between international bodies, different official publications of the U.S. government, and among individuals who potentially might identify with this group.

It seems futile to attempt to resolve this disagreement for the purposes of federal minimum standards around race and ethnicity data. For our purposes, it is not necessary to rehearse the historically fraught colonial project of defining boundaries and characterizing the Middle East distinct from other parts of the Asian continent or the separation of North Africa from Africa south of the Sahara. Indeed, the treatment of the globe as consistent of seven distinct continents is a construction arising out of the needs of the imperialist project (Kaplan, 2024).

[6] An alternative that has been adopted by some in the community is SWANA or South West Asia and North Africa.

What is clear is that even if we were to resolve what countries constitute the Middle East and North Africa it would still not make sense to offer a racial category in this way for the purposes of Directive 15 because racialization is not primarily about the geographical origin.

Discussions around adding a MENA category, perhaps alongside others that allows respondents to select Black or African American, White, or something else in addition to MENA, reveal a similar dynamic as the Hispanic group. People are racialized as Arab/Muslim or as Hispanic/Mexican/Latino based on phenotype, cultural markers, language use, and local context, but the definitions do not correspond to the ways in which individuals locate themselves across the categories, instead focusing on having "origins" in the "original peoples" of various ambiguously defined geographical regions.

These definitions simply are not necessary. Each of our critiques taken individually seem to suggest the possibility of an alternative definition that is not subject to the critique. For example, in this case we could rewrite the definitions to remove or deemphasize references to "original peoples" from various continents.

But, taken together, the critiques imply that it is not possible to generate usable, practical definitions for the intended purposes of Directive 15, much less move these forward given the political and administrative realities that are documented in the record of the inter-agency deliberation and public consultation for the 1977 drafting and the 1997 and 2024 revisions. The fact that the standards were written originally and revised with many of these problems acknowledged indicates that the political process of achieving consensus between different civil society constituencies and stakeholders within the federal government is difficult, costly, and always results in compromise.

To be clear, what makes the Directive 15 definitions wrong is not just that they are built on debunked racial science, but that they function to naturalize a social phenomenon. Critical black thought in the U.S. has, since the beginnings of white supremacy (Robinson & Kelley, 2000), recognized that what separates whites from blacks in the U.S. is not "having origins in the original peoples of" different regions, but group-based differences in power.

Race is invented, and racism is maintained, challenged, and circumvented through historical struggle. Racism endures because it accomplishes instrumental aims, namely, material and psychological benefits to members of the dominant group at the expense of members of the marginalized group (Mason, 2023). Racial groups do not need to be "defined" for this system to operate.

Toro (1995, p. 1244) articulates a widely accepted view in the social sciences applied specifically to Directive 15:

> The drafters of Directive No. 15 did not attend to this crucial distinction between the fiction of race and the reality of racism. Instead, they assumed that a person can be said to belong to a race, and that racism occurs when someone falsely attributes other characteristics to a person that are dependent on the physical, racial appearance of that person.

Efforts to denaturalize racial difference interrupt the operation of symbolic violence (Swartz, 1997), the process by which social group-based inequality is justified and becomes part of dominant, hegemonic understandings of the world. Removing the definitions from Directive 15 constitutes one more step towards denaturalizing racial categories.

3.3 Inappropriately constraining identity

Directive 15 has taken on a life of its own, beyond the scope of its stated purpose. The administrative practices of the nation-state influence peoples' identities and the realities of group-based inequality. Initially designed to collect data, Directive 15 has played a role in how Americans view themselves.

As we contemplate updates to Directive 15 in the future, we must reckon with the Directive's role in reifying particular understandings of racial identities. Walter and Anderson (2013, p. 8) write from the standpoint of indigenous scholarship: "As we invest ourselves and our communities in their categories, we increasingly use statistics to help us tell *ourselves* who we are." Or, as the late Pierre Bourdieu (2009, p. 224) wrote, "objectivist arbitration... can give rise in reality, by specific effectiveness of evocation, to the very thing they represent."

The categories and definitions in Directive 15 are meant to reflect social reality, but they have the effect of reflexively affecting our social reality. To the extent that the definitions reflect dominant ideas about race, they may shape social reality in ways that maintain or deepen racial hierarchies rather than challenge them.

The current definitions ascribe attributes to groups that may not be embraced by all persons who identify with the group, and they can omit attributes that some members would find indispensable to their understanding of their identity. These distortions need to be balanced with the potential

benefits that can come from federally sponsored data collection. We find that, in light of the proposed change to Directive 15, this balance can be better achieved by providing category names without definitions.

There are many examples of inappropriate constrains on identity choices that would not be present absent the definitions. Persons choosing the black category, particularly if they are descendants of persons enslaved in the U.S., may self-identify as black without the definition of their status referring to an African country of origin, particularly af they do not necessarily want to do so.

With respect to ethnic identification, these individuals may see themselves as descendants of American Freedmen, in contrast with a smaller share of the self-reported black population comprised of more recent immigrants to the U.S. or the much smaller share who are descendants of the American Indian Freedmen emancipated from the "Five Civilized Tribes." Only the second group may possess a primary sense of "origin" in and identification with countries in Africa or Africa writ large.

Under the current system, in principle, anyone with any known white ancestry should identify as white. People routinely ignore this, which is both a reflection of historical hypodecendency but also a personal choice to identify with a different community. These choices can be strategic, reflecting individual utility maximization (Darity, Mason, & Stewart, 2006), they may be ethical positions, the result of sociocultural phenomena, or some combination.

In either case, including the definitions in Directive 15 injects needless confusion into the effectiveness of identification choices: individuals are free to believe what they like under the current standards, but are limited in their ability to express carefully considered positions for the purposes of government and non-government statistics.

4 Proposal to delete racial group definitions in Directive 15

The benefits of federal minimum standards for the collection and reporting of data on race and ethnicity are substantial. We find that this proposal for deleting the definitions is compatible with many different future revisions, and it will better enable achievement of Directive 15's aims. In many cases, this proposal would not result in any change in actual practice because the definitions are

rarely used explicitly, but it does represent a profound symbolic change and it resolves contradictions in how the policy works in practice.

Under the proposed change, the federal government would set out a minimum set of categories and provide other guidance around data collection and reporting. Individuals might be asked if they identify as or are part of one of the groups, say, white, but this term would not be defined. In practice, surveys rarely provide the definitions from Directive 15, so this would imply little change in most cases.

Asking individuals to identify with categories while not providing a definition is a common practice. This is typically how questions are asked concerning gender. As with race, there are many different ways to solicit this information, and the number and meaning of categories can vary from survey to survey.

The number of categories and which specific ones to include is—and rightly should be—a key matter of debate in each round of revision of the standards. The original 1977 policy had four racial categories, the revised 1997 version had five racial categories, and the latest issued in March 2024 has seven race and/or ethnicity categories.

Individuals with strong preferences about adding particular categories can support the proposed change since it would increase the freedom of individuals and groups to define and implicitly assign characteristics to categories. The possibility of deleting caategories also always will be available. The proposal does not require or preclude changes to the number or arrangement of the minimum standard categories.

The proposal under consideration will eliminate the use of definitions in the required minimum standards. Each round of revision is accompanied by an extensive consultation period with stakeholders both outside and within the federal government, including other organizations that share data with the federal government, community groups and advocates. While it would be tempting to write from scratch our preferred Directive 15 text, we value the democratic potential of deliberation that is part of the revision process and also the specialized knowledge that federal administrators have around the costs of implementing changes to Directive 15 in terms of survey length, data processing and storage, and redesigning surveys. The proposal is narrowly tailored and consistent with these considerations for the administrative process of revising Directive 15.

5 Policy evaluation and discussion

5.1 Evaluating the proposal

In this section, we evaluate the proposal to remove the Directive 15 definitions that define membership in racial groups. Obviously, all the problems discussed in Section 3 of this paper are resolved immediately, eradicating (1) the inconsistency, noncomprehensiveness, and circularity of definitions, (2) the problems with defining race in a way that conflates race with other factors such as nationality or refers to the so-called "original peoples" of various continents, and (3) the difficulties with the federal government inappropriately defining membership and characteristics of racial groups.

Eliminating the definitions from the minimum standards also removes a significant piece of controversy around the federal handling of data on race, while yet preserving comparability and flexibility in how data is collected and reported.

Another justification for eliminating the definitions concerns white people's understanding of their own racial identity as key to fostering a commitment to take anti-racist action (Wingfield, 2015). That is, individuals interrogating how their own race affects their life is consciousness raising and, for white Americans, this has little to do with having "origins in any of the original peoples of Europe," *per se*. The definitions in Directive 15 constitute an act of official messaging by the federal government defining the boundaries and characterizing the broad racial groups to which, presumably, we all belong. Even if the benefits from the proposed change are "merely" symbolic, race operates significantly on the level of representation, and ideology is a necessary conjugate to maintain its instrumental purpose of unfairly allocating material resources and life chances.

The proposed change to Directive 15 is also justified with reference to the modern understanding of race as a sociopolitical construct that has the appearance of being an individual quality but is more accurately understood as primarily a group-based phenomenon. Some people experience their racial identity as context-specific and multifaceted, and they often experience conflict between self-identity and outward perceptions.

The federal definitions give the appearance of creating a straightforward criterion for racial identity, but the contradictions of race are not and cannot be resolved in the way that Directive 15 approaches racial definitions. Racism as a system is resistant to change not because of its coherence

but, in some ways, because of its contradictions. M. Omi (1985, p. 21) writes:

> Among scholars there is a continuous temptation to think of race as an essence, as something fixed, concrete and objective. There is also an opposite temptation: to imagine race as a mere illusion, a purely ideological construct which masks some other more fundamental division, such as class.

As M. Omi (1985) argues, both of these approaches are inadequate.

Data on race is important primarily because it is necessary for quantifying how hierarchy predicated on social identity groups manifests in society and for understanding how we might better pursue change towards greater racial justice and liberation. We can balance the benefits of data collection, which provides a powerful way of understanding the reality of racial hierarchy as a produced socio-political phenomenon, with the symbolic violence of asking people to reduce this complexity to a choice made from a list of pre-populated categories by having the federal government not dictate how the groups are characterized.

5.2 Practical considerations

How would this proposed change work in practice? In this section, we discuss what we imagine statistical practices would look like if definitions were removed from Directive 15. If the proposal were implemented, the government would collect data on surveys, administrative forms, or other methods, much as it currently does. Neither the phrasing of the questions on a Census or American Community Survey nor the categories offered would necessarily change under this proposal.

The existing definitions provide a false sense of objectivity and knowledge. They appear to tell us, definitively, what it means when someone selects a particular racial group (i.e. that they have "origins in any of the original peoples..." of a particular continent), but this is in fact inadequate for understanding an individual's selection(s). It may be besides the point.

Removal of the definitions would mean revisiting Directive 15 in the future would not require revising the definitions. It will require potentially revising the categories themselves for consistency with extant social usage, but this makes complete sense, since the meaning of race is socially constructed and changing over time.

In summary, the proposed change would end the regular need to update definitions, the regular

political fights about the definitions, and it would end on-going symbolic violence associated with the definitions.

5.3 Clarifying what the proposal is not

There are political efforts aimed at curtailing critical attention to racial inequality and one of their manifestations is advocacy around statistical data collection. Many different proposals have been structured with the instrumental intent of reducing critical attention to racial hierarchies (Prewitt, 2016). M. A. Omi (2001, p. 248) cites one example, writing, "House Speaker Newt Gingrich... used the issue of multiraciality to illustrate the indeterminacy of racial categories and to vigorously advocate for their abolition in government data collection, much as advocates of color-blindness do."

In contrast, this proposed change is consistent with more accurate collection and reporting around race. It is consistent with calls for more disaggregated data and calls for attention to the multidimensional nature of race, including more data collected on skin shade and other phenotypical characteristics that may dictate social identification of an individual's race (Diette, Goldsmith, Hamilton, & Darity, 2015; Goldsmith, Hamilton, & Darity, 2006; Goldsmith, Hamilton, & Darity Jr., 2007; Hamilton, Goldsmith, & Darity, 2009; Jones, 2000).

Going further, some have argued that we should pursue a strategy that minimizes or eliminates any explicit mention of race as a precondition of anti-subordination (A. M. Johnson, 1996) or that multiracial identities have the potential to destabilize racial hierarchies. We disagree. As Kimberlé Crenshaw (1988, p. 1336) writes, "History has shown that the most valuable political asset of the Black community has been its ability to assert a collective identity and to name its collective political reality."

Regardless, removing the definitions from Directive 15 does not require or incentivize the government to collect less (or more) data on race and ethnicity. Whether one believes that anti-racism requires abandoning racial identity or, conversely, that it requires continued political organization along racial identity lines, one can support the idea of removing definitions from Directive 15.

5.4 Comparability of data on race over time

Finally, a comment about comparability of data over time. At the level of data collection, holding aside immigration, demographics, and related issues, the identity of any particular racial group obtained through surveys changes over time due to two factors: (1) changes in the social meaning attached to different groups or terms and (2) changes in survey methodology.

For example, individuals with both white and black parentage changed their behavior in surveys coincident with the election and visibility of President Barack Obama, who identified strongly as someone with both black and white heritage (Mason, 2017). This form of change over time reflects the fundamentally social nature of race. As another example, "Prior to the late 1960s. . . there were no people who identified as 'Asian American"' because that category did not exist (Espiritu, 1993, p. 17).

But there are also changes in the data that come from purely methodological decisions. The composition of Asian or Pacific Islander changed when the group's definition changed in 1997 with a separate category for Native Hawaiians and Other Pacific Islanders.

The meaning of race as reflected in data sets changes over time and the proposal in question would eliminate one source of variation: changes to the definitions. The problem of keeping up with social categories remains. This problem has additional complexity in the context of the federal government because, "as racial categories change, so do racial hierarchies" (Bashi, 1998). Visible decisions around statistical practices by the federal government, including Directive 15 and the decennial Census, have the power to affect social reality even when their primary intended purpose is to facilitate quantitative measurement of social reality.

Seen in this light, perhaps the observation that "no single set of racial categories has been used in more than two censuses, and most were only used once" (Martin et al., 1990, p. 553) is rather hopeful, compared to the rigidity of the definitions of Directive 15 which have remained largely unchanged since 1977.

6 Conclusion

The race categories identified by Census and OMB have powerful social meaning. Self-reported race correlates strongly with a wide range of social outcomes. But self-reported race is not selected

based on the official definitions.

Given the inadequacy of the definitions and the fact they do not seem to inform the choices made by respondents to federal surveys, we conclude that dispensing with the definitions altogether could have substantial advantages over the status quo routine of intermittently trying to update specific definitions for race and ethnicity categories. Independent research and opportunities to receive public comments and recommendations can ensure the Office of Management and Budget and the U.S. Census Bureau are making use of an appropriate list of categories that are popularly understood and socially meaningful.

References

94th Congress. (1976, June). *Joint resolution relating to the publication of economic and social statistics for Americans of Spanish origin or descent.*

American Anthropological Association. (1998). *AAA Statement on Race.* Retrieved from `https://americananthro.org/about/policies/statement-on-race`

American Association of Biological Anthropologists. (1996). *AABA Statement on Race.* Retrieved from `https://bioanth.org/about/aaba-statement-on-race`

Bashi, V. (1998, January). Racial categories matter because racial hierarchies matter: a commentary. *Ethnic and Racial Studies*, *21*(5), 959–968. doi: 10.1080/014198798329748

Bourdieu, P. (2009). *Language and symbolic power* (J. B. Thompson & G. Raymond, Trans.). Cambridge: Polity Press.

Citro, C. F. (1997, September). Window on Washington: Recognizing Diversity: Recommendations to OMB on Standards for Data on Race and Ethnicity. *CHANCE*, *10*(4), 26–31. doi: 10.1080/09332480.1997.10542059

Crenshaw, K. W. (1988). Race, Reform, and Retrenchment: Transformation and Legitimation in Antidiscrimination Law. *Harvard Law Review*, *101*(7), 1331–1387.

Darity, W. A., Mason, P. L., & Stewart, J. B. (2006, July). The economics of identity: The origin and persistence of racial identity norms. *Journal of Economic Behavior & Organization*, *60*(3), 283–305. doi: 10.1016/j.jebo.2004.09.005

Darity, W. A., & Mullen, A. K. (2020). *From Here to Equality: Reparations for Black Americans in the Twenty-First Century.* Chapel Hill: The University of North Carolina Press.

Desautels-Stein, J. (2012). Race as a Legal Concept. *Columbia Journal of Race and Law*, *2*(1), 1–74.

Diette, T. M., Goldsmith, A. H., Hamilton, D., & Darity, W. (2015, January). Skin Shade Stratification and the Psychological Cost of Unemployment: Is there a Gradient for Black

Females? *The Review of Black Political Economy*, *42*(1-2), 155–177. doi: 10.1007/s12114-014-9192-z

Eschbach, K., & Taylor, J. (in press). The Explosion of the 2020 Census "Indian" Population and Its Implications for Tribal Policymaking. In N. Hill, K. Ratteree, M. Hill, & D. Hill (Eds.), *Beyond Blood Quantum: Refusal to Disappear.* Fulcrum Publishing.

Espiritu, Y. (1993). *Asian American Panethnicity: Bridging Institutions and Identities.* Philadelphia: Temple University Press.

Fears, D. (2003, October). The Roots of 'Hispanic'; 1975 Committee of Bureaucrats Produced Designation. *The Washington Post*. (Oct 15, 2003)

Federal Interagency Committee on Education. (1975, April). *Report of the Ad Hoc Committee on Racial and Ethnic Definitions.*

Fields, B. J. (1982, June). Ideology and Race in American History. In J. M. Kousser & J. M. McPherson (Eds.), *Region, Race and Reconstruction: Essays in Honor of C. Vann Woodward* (pp. 143–177). New York: Oxford University Press.

Forbes, J. D. (1992, October). The Hispanic Spin: Party Politics and Governmental Manipulation of Ethnic Identity. *Latin American Perspectives*, *19*(4), 59–78. doi: 10.1177/0094582X9201900406

Ford, C. A. (1994, October). Administering Identity: The Determination of "Race" in Race-Conscious Law. *California Law Review*, *82*(5), 1231. doi: 10.2307/3480910

Fuentes, A., Ackermann, R. R., Athreya, S., Bolnick, D., Lasisi, T., Lee, S.-H., ... Nelson, R. (2019). AAPA Statement on Race and Racism. *American Journal of Physical Anthropology*, *169*(3), 400–402. doi: 10.1002/ajpa.23882

Gimenez, M. E. (1989, July). Latino/"Hispanic"—Who Needs a Name? The Case against a Standardized Terminology. *International Journal of Health Services*, *19*(3), 557–571. doi: 10.2190/HN6N-P1TH-8CHL-KW5X

Goldsmith, A. H., Hamilton, D., & Darity, W. (2006, April). Shades of Discrimination: Skin Tone and Wages. *American Economic Review*, *96*(2), 242–245. doi: 10.1257/000282806777212152

Goldsmith, A. H., Hamilton, D., & Darity Jr., W. A. (2007). From Dark to Light: Skin Color and Wages Among African-Americans. *Journal of Human Resources*, *XLII*(4), 701–738. doi: 10.3368/jhr.XLII.4.701

Hamilton, D., Goldsmith, A. H., & Darity, W. (2009, October). Shedding "light" on marriage: The influence of skin shade on marriage for black females. *Journal of Economic Behavior & Organization*, *72*(1), 30–50. doi: 10.1016/j.jebo.2009.05.024

Hernández, T. K. (1998). Multiracial Discourse: Racial Classifications in an Era of Color-blind Jurisprudence. *Maryland Law Review*, *57*, 97–173.

Hitlin, S., Brown, J. S., & Elder, G. H., Jr. (2007, December). Measuring Latinos: Racial vs.

Ethnic Classification and Self-Understandings. *Social Forces*, *86*(2), 587–611. doi: 10.1093/sf/86.2.587

Hollinger, D. A. (2000). *Postethnic America: beyond multiculturalism.* New York: BasicBooks.

Johnson, A. M. (1996). Destabilizing Racial Classifications Based on Insights Gleaned from Trademark Law. *California Law Review*, *84*(4), 887–952. (Publisher: California Law Review, Inc.) doi: 10.2307/3480986

Johnson, A. M., Jr. (2008). The Re-Emergence of Race as a Biological Category: The Societal Implications—Reaffirmation of Race. *Iowa Law Review*, *94*(1547-1585).

Jones, T. (2000). Shades of Brown: The Law of Skin Color. *Duke Law Journal*, *49*.

Larena, M., Sanchez-Quinto, F., Sjödin, P., McKenna, J., Ebeo, C., Reyes, R., ... Jakobsson, M. (2021, March). Multiple migrations to the Philippines during the last 50,000 years. *Proceedings of the National Academy of Sciences*, *118*(13), e2026132118. doi: 10.1073/pnas.2026132118

Lew, J. J. (2000, March). *Guidance on Aggregation and Allocation of Data on Race for Use in Civil Rights Monitoring and Enforcement. Office of Management and Budget.*

López, N., & Hogan, H. (2021, August). What's Your Street Race? The Urgency of Critical Race Theory and Intersectionality as Lenses for Revising the U.S. Office of Management and Budget Guidelines, Census and Administrative Data in Latinx Communities and Beyond. *Genealogy*, *5*(3), 75. doi: 10.3390/genealogy5030075

Margolis, M. (1995, March). Brazilians and the 1990 United States Census: Immigrants, Ethnicity, and the Undercount. *Human Organization*, *54*(1), 52–59. doi: 10.17730/humo.54.1.5487w18130176194

Martin, E., Demaio, T. J., & Campanelli, P. C. (1990). Context Effects for Census Measures of Race and Hispanic Origin. *Public Opinion Quarterly*, *54*(4), 551. doi: 10.1086/269227

Martinez, G. A. (1997). The Legal Construction of Race: Mexican-Americans and Whiteness. *Harvard Latino Law Review*, *2*, 321–347.

Mason, P. L. (2017, January). Not Black-Alone: The 2008 Presidential Election and Racial Self-Identification among African Americans. *The Review of Black Political Economy*, *44*(1-2), 55–76. doi: 10.1007/s12114-017-9247-z

Mason, P. L. (2023). *The Economics of Structural Racism.* Cambridge, United Kingdom ; New York, NY: Cambridge University Press.

Mathews, K., Phelan, J., Jones, N. A., Konya, S., Marks, R., Pratt, B. M., ... Bentley, M. (2017, February). *2015 National Content Test: Race and Ethnicity Analysis Report: A New Design for the 21st Century.*

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Llamas, B., Pickrell, J., ... Reich, D. (2015, March). *Eight thousand years of natural selection in Europe.* bioRxiv. (Pages: 016477 Section: New Results) doi: 10.1101/016477

Mirandé, A. (Ed.). (2019). *Gringo Injustice: Insider Perspectives on Police, Gangs, and Law*. New York London: Routledge.

Mora, G. C. (2014). *Making Hispanics: How Activists, Bureaucrats, and Media Constructed a New American*. Chicago ; London: University of Chicago Press.

Myers Jr., S. L. (1997). The Error of the Third Type. In C. W. Hartman (Ed.), *Double Exposure: Poverty and Race in America.* New York: Routledge. doi: 10.4324/9781315479378

Office of Management and Budget. (1978, May). *Transfer of Responsibility for Certain Statistical Standards from OMB to Commerce.* (Fed. Reg. Vol. 43, No. 87)

Office of Management and Budget. (1994, June). *Standards for the Classification of Federal Data on Race and Ethnicity.* (Fed. Reg. Vol. 59, No. 110)

Office of Management and Budget. (1995, August). *Standards for the Classification of Federal Data on Race and Ethnicity.* (Fed. Reg. Vol. 60, No. 166)

Office of Management and Budget. (1997, October). *Revisions to the Standards for the Classification.* (Fed. Reg. Vol. 62, No. 210)

Office of Management and Budget. (2023, January). *Initial Proposals For Updating OMB's Race and Ethnicity Statistical Standards.*

Office of Management and Budget. (2024, March). *Revisions to OMB's Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity.* (Fed. Reg. Vol. 89, No. 62)

Omi, M. (1985). Racial Identity and the State: The Dilemmas of Classification. *Minnesota Journal of Law & Inequality*, *15*(1).

Omi, M. A. (2001, January). The Changing Meaning of Race. In N. J. Smelser, W. J. Wilson, & F. Mitchell (Eds.), *America Becoming: Racial Trends and Their Consequences* (Vol. 1, pp. 243–263). Washington, D.C.: National Academies Press. (Pages: 9599) doi: 10.17226/9599

Prewitt, K. (2016). *What Is "Your" Race?: The Census and Our Flawed Efforts to Classify Americans*. Princeton University Press.

Ray, V. (2019). A Theory of Racialized Organizations. *American Sociological Review*, *84*(1), 28.

Robbin, A. (2000, July). The politics of representation in the US national statistical system: origins of minority population interest group participation. *Journal of Government Information*, *27*(4), 431–453. doi: 10.1016/S1352-0237(00)00178-7

Robinson, C. J., & Kelley, R. D. G. (2000). *Black Marxism: The Making of the Black Radical Tradition*. Chapel Hill, N.C: The University of North Carolina Press.

Rodríguez, C. E. (1997). *Latin Looks: Images Of Latinas And Latinos In The U.S. Media*. Boulder, Colo: Routledge.

Rosa, J. (2019). *Looking Like a Language, Sounding Like a Race*. Oxford University Press.

Rosa, J., & Flores, N. (2017, November). Unsettling race and language: Toward a raciolinguistic perspective. *Language in Society*, *46*(5), 621–647. (Publisher: Cambridge University Press) doi: 10.1017/S0047404517000562

Roth, W. D. (2016, June). The multiple dimensions of race. *Ethnic and Racial Studies*, *39*(8), 1310–1338. doi: 10.1080/01419870.2016.1140793

Schuster, L. (2021, December). We're Reporting Census Data All Wrong. *Boston Indicators*. Retrieved 2024-02-27, from `https://www.bostonindicators.org/article-pages/2021/december/census_reporting`

Shadowen, S. D. (2018). U.S. Border Patrol's Policy of Extrajudicial Killing. *Berkeley La Raza Law Journal*, *28*(1).

Skerry, P. (2000). *Counting on the Census?: Race, Group Identity, and the Evasion of Politics*. Washington, D.C: Brookings Institution Press.

State Health Access Data Assistance Center at the University of Minnesota. (2021, May). *Collection of Race, Ethnicity, Language (REL) Data in Medicaid Applications: A 50-state Review of the Current Landscape.*

Swartz, D. (1997). *Culture and Power: The Sociology of Pierre Bourdieu*. University of Chicago Press.

Tamayo Lott, J. (1997). The Limitations of Directive 15. In *Double Exposure: Poverty and Race in America.* New York: Routledge. doi: 10.4324/9781315479378

Toro, L. A. (1995). A People Distinct from Others: Race and Identity in Federal Indian Law and the Hispanic Classifications in OMB Directive No. 15. *Texas Law Review*, *26*(4), 1219–1274.

Wallman, K. K. (1998, February). Data on Race and Ethnicity: Revising the Federal Standard. *The American Statistician*, *52*(1), 31–33. doi: 10.1080/00031305.1998.10480533

Wallman, K. K., Evinger, S., & Schechter, S. (2000, November). Measuring Our Nation's Diversity: Developing a Common Language for Data on Race/Ethnicity. *American Journal of Public Health*, *90*(11), 1704–1708.

Walter, M., & Anderson, C. (2013). *Indigenous statistics: a quantitative research methodology*. Walnut Creek, CA: Left Coast Press.

Wingfield, A. H. (2015, September). Color Blindness Is Counterproductive. *The Atlantic*. Retrieved 2024-03-05, from `https://www.theatlantic.com/politics/archive/2015/09/color-blindness-is-counterproductive/405037/`

Yang, T. (2005). Choice and Fraud in Racial Identification: The Dilemma of Policing Race in Affirmative Action, the Census, and a Color-Blind Society. *Michigan Journal of Race & Law*, *11*.

Table 1. OMB Statistical Directive No. 15 definitions: Original 1977 policy and updates made in 1997 and 2024

Category	Definition 1977	Category	Definition 1997	Category	Definition 2024
American Indian or Alaskan Native	A person having origins in any of the original peoples of North America, and who maintains cultural identification through tribal affiliation or community recognition.	No change	A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.	No change	Individuals with origins in any of the original peoples of North, Central, and South America, including, for example, Navajo Nation, Blackfeet Tribe of the Blackfeet Indian Reservation of Montana, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, Aztec, and Maya.
Asian or Pacific Islander	A person having origins in any of the original peoples of the Far East, Southeast Asian, the Indian subcontinent, or the Pacific Islands. This area includes, for example, China, India, Japan, Korea, the Philippine Islands, and Samoa.	Asian	A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.	No change	Individuals with origins in any of the original peoples of Central or East Asia, Southeast Asia, or South Asia, including, for example, Chinese, Asian Indian, Filipino, Vietnamese, Korean, and Japanese.
Black	A person having origins in any of the black racial groups of Africa.	Black or African American	A person having origins in any of the black racial groups of Africa. Terms such as "Haitian" or "Negro" can be used in addition to "Black or African American."	No change	Individuals with origins in any of the Black racial groups of Africa, including, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, and Somali.
Hispanic	A person or Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or origin, regardless of race.	Hispanic or Latino	A person of Cuban, Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race. The term, "Spanish origin," can be used in addition to "Hispanic or Latino."	No change	Includes individuals of Mexican, Puerto Rican, Salvadoran, Cuban, Dominican, Guatemalan, and other Central or South American or Spanish culture or origin.
White	A person having origins in any of the original peoples of Europe, North Africa, or the Middle East.	No change	A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.	No change	Individuals with origins in any of the original peoples of Europe, including, for example, English, German, Irish, Italian, Polish, and Scottish.
		Native Hawaiian or Other Pacific Islander	A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.	Native Hawaiian or Pacific Islander	Individuals with origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands, including, for example, Native Hawaiian, Samoan, Chamorro, Tongan, Fijian, and Marshallese.
				Middle Eastern or North African	Individuals with origins in any of the original peoples of the Middle East or North Africa, including, for example, Lebanese, Iranian, Egyptian, Syrian, Iraqi, and Israeli.