



# POLICY BRIEF

## The Census Bureau's Proposed "Combined Question" Approach Offers Promise for Collecting More Accurate Data on Hispanic Origin and Race, but Some Questions Remain

### The Importance of Accurate Census Data for the Latino Community

The accuracy of the data collected by the Census Bureau on the nation's population, and on its racial, ethnic and national origin groups, helps ensure our nation's future prosperity and well-being. These data guide a wide range of decisions made in the public and private sectors that affect the lives of Latinos and all American families and their children. The data help ensure fair and representative reapportionment and redistricting for Latino communities. Census data play an indispensable role in the monitoring and implementation of civil rights policies, and they are used to allocate billions of dollars in federal, state and local funding.

Latinos are the nation's second largest population group, and one of its fastest growing communities. Latinos account for more than one of every six U.S. residents, and one of every four of the country's population under 18. For the Census Bureau to compile the most accurate data possible about the U.S. population, it must ensure a full and accurate count of the Latino community.

For Census data to present an accurate portrait of our Latino population, they must reflect the on-going evolution of Americans' racial and ethnic identity. The Census Bureau has undertaken a comprehensive evaluation to determine whether changes to the wording and format of its questions on Hispanic origin and race would improve the accuracy of the responses it receives.

#### **The Office of Management and Budget Standards for Federal Data on Race and Ethnicity:**

The Census Bureau collects data on race and ethnicity in accordance with standards for federal data established by the Office of Management and Budget (OMB). Federal data on race and ethnicity serve two crucial purposes that are particularly important for the Latino community. The data reveal persistent disparities rooted in historic discrimination premised on race and ethnicity, and they enable efforts to eliminate those disparities.

In 1997, the OMB undertook a major revision of its classification standards, and adopted an approach which defined two ethnicity categories - Latino, or not Latino. The standards required that race be measured separately from these ethnicity categories, and defined five minimum race categories - generally, White, Black, American Indian/Alaska Native, Asian, or Native Hawaiian/Pacific Islander.

# The Census “Two Separate Question” Approach to Hispanic Origin and Race

Beginning in 1970, the Census included separate questions on Hispanic identity and racial identity in its questionnaires. Research in the 1990s showed that the government derived much better data with two separate questions than a combined question. In addition, the Census questions on Hispanic origin and race continued to evolve after the OMB revised its standards and established two ethnicity categories and five minimum race categories.

The 2010 Census questions reflect the “two separate question” approach to obtaining responses on Hispanic origin and race. The questionnaire first asked individuals to indicate whether or not they were of Hispanic origin, followed by a separate question asking respondents to indicate their race. Respondents were provided five general race categories: White; Black; American Indian or Alaska Native; a selection of nine specific large Asian and Native Hawaiian or Pacific Islander subgroups, with two additional write-in boxes for “Other Asian” and “Other Pacific Islander; and “Some other race.” Respondents could also select one or more race (Figure 1)

Figure 1

→ NOTE: Please answer BOTH Question 8 about Hispanic origin and Question 9 about race. For this census, Hispanic origins are not races.

**8. Is Person 1 of Hispanic, Latino, or Spanish origin?**

No, not of Hispanic, Latino, or Spanish origin

Yes, Mexican, Mexican Am., Chicano

Yes, Puerto Rican

Yes, Cuban

Yes, another Hispanic, Latino, or Spanish origin — *Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.* ↴

**9. What is Person 1's race? Mark  one or more boxes.**

White

Black, African Am., or Negro

American Indian or Alaska Native — *Print name of enrolled or principal tribe.* ↴

<input type="checkbox"/> Asian Indian	<input type="checkbox"/> Japanese	<input type="checkbox"/> Native Hawaiian
<input type="checkbox"/> Chinese	<input type="checkbox"/> Korean	<input type="checkbox"/> Guamanian or Chamorro
<input type="checkbox"/> Filipino	<input type="checkbox"/> Vietnamese	<input type="checkbox"/> Samoan
<input type="checkbox"/> Other Asian — <i>Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.</i> ↴	<input type="checkbox"/> Other Pacific Islander — <i>Print race, for example, Fijian, Tongan, and so on.</i> ↴	

Some other race — *Print race.* ↴

→ If more people were counted in Question 1, continue with Person 2.

## Reasons for the Census Bureau's Redesign of Hispanic Origin and Race Questions

### **The mismatch between Latino identification and the Census race categories:**

As part of its effort to increase the reliability of responses and to increase census response rates, the Census Bureau continuously examines the format and wording of the Hispanic origin and race questions. The Bureau has long recognized a growing mismatch between the specific categories set forth in the race question and the ways in which many Latinos express their racial and ethnic background. The Bureau has consistently found that Latinos account for majorities of people who do not report themselves as belonging to any of the specific race categories. For example, more than 43% of Latinos chose "Some other race" or did not answer the race question on the 2010 decennial Census. Of those who chose "Some other race" and wrote in the race with which they identified, an overwhelming majority answered "Mexican," "Hispanic," "Latin American," or "Puerto Rican," which suggests they did not identify with the specific racial groups set forth in the question.

### **Consistency between Census data and OMB categories:**

The relatively high number of Latinos who do not identify with the specific racial categories set forth in the Census question also presents a major challenge for the consistency of Census data with race and ethnic categories established by the OMB. Because the OMB minimum race categories do not include "Some other race," the Bureau has developed a procedure to assign an OMB race to those checking "Some other race" in response to the Census question. The procedure relies on assessment of the demographic characteristics of these respondents and their family members and neighbors, and the Bureau uses it for post-Census calculations and products, such as population estimates.

Between 2000 and 2010, the population classified as "Some other race" grew by one-quarter. By 2010, 6% of all decennial Census respondents – 19.1 million people – identified themselves as "Some other race," and 97% of those 19.1 million individuals were Latino. Latinos aged 18 to 44 are statistically more likely than their older counterparts to answer "Some other race" or to provide no answer to a separate Census question about race. The Census Bureau expects that by 2020, "Some other race" could become the second largest racial group reported in the Census.

### **Missing information about detailed national origin:**

Finally, the Census Bureau's research around alternate Hispanic origin and race question design has been motivated by the goal of collecting more accurate detailed data about Latino national origin and sub-groups. The separate-question format employed in the 2010 decennial Census did not accommodate the reporting of multiple national origins by respondents of Latino ethnicity. However, evidence indicates that inability to record and report multiple Hispanic national origins results in the loss of detailed information about Latino identities. For example, in response to a differently-formatted ethnicity question on the 2000 decennial Census, the Bureau observed that more than 260,000 respondents attempted to report multiple Latino national origins, and that such reporting was most common among respondents under the age of 35, portending future increases in the percentages of Latinos identifying as being of more than one Latino national origin.

## The Census Bureau's 2010 Research on the Hispanic Origin and Race Questions

The Census Bureau's research to enhance its collection of race and ethnicity data includes its Alternative Questionnaire Experiment (AQE), which was conducted during Census 2010 to test a number of different questionnaire design strategies. During the 2010 enumeration, the Bureau mailed out questionnaires with several different experimental variations in the design strategies for the Hispanic origin and race questions.

The Bureau also conducted interviews and focus groups with respondents, which revealed that some Latinos struggled to answer a separate question about race, and were not necessarily satisfied with the answers they had given. This research indicated that a number of Latinos reported their race as White to the separate race question because they did not see themselves represented among the other options in the race question and felt forced to choose a race category. The Bureau found that significant numbers of Latinos did not embrace or express any identity other than that stemming from their Latino and specific national origin heritage – in other words, Latino was equivalent to these individuals' "race." However, the Bureau also found that some survey respondents strongly resist the equating of nationality or ethnicity to race, and insisted on the importance of distinguishing between geographic origins and pan-national racial classifications.

In addition, the Census Bureau's research raised concerns about how non-Hispanic individuals responded to the questions on Hispanic origin and race. For example, some non-Hispanics did not respond to the separate Hispanic origin question, because they did not feel it applied to them or believed they had answered the question when responding to the question on race.

Survey participants also told the Census Bureau that they were skeptical or suspicious of the motives for asking a separate question that sought only to identify people with Latino ethnicity, for various reasons. Some individuals worried that the question enabled the discriminatory treatment of people who denoted themselves as Hispanic or Latino. At the same time, some members of the public saw the question as unfair because it could lead to provision of special, preferential treatment to Latinos. In either case, survey subjects perceived Latinos being counted differently than other groups of Americans because Latino ethnicity was the sole topic of a dedicated question.

# The Census Bureau's 2015 Research on the Combined Race and Ethnicity Question Approach

In addition to the 2010 AQE, the Bureau undertook a comprehensive 2015 evaluation of the content of the Census questionnaire - the 2015 National Content Test (NCT). The 2010 AQE and the 2015 NCT sought to determine what question format would achieve the most in terms of decreasing nonresponse, increasing reporting within the OMB-recognized race and ethnicity categories, eliciting detailed reporting of Latino national origin and sub-groups, and increasing accuracy and reliability of results.

The Census Bureau's research appears to demonstrate that a single combined question on race and ethnicity accompanied by detailed checkboxes best achieves the foregoing goals (Figure 2).

Figure 2

**8. Which categories describe Person 1?**  
 Mark all boxes that apply **AND** print details in the spaces below.  
 Note, you may report more than one group.

**WHITE** – Provide details below.

German       Irish       English  
 Italian       Polish       French

Print, for example, Scottish, Norwegian, Dutch, etc. ⌘

**HISPANIC, LATINO, OR SPANISH** – Provide details below.

Mexican or Mexican American       Puerto Rican       Cuban  
 Salvadoran       Dominican       Colombian

Print, for example, Guatemalan, Spaniard, Ecuadorian, etc. ⌘

**BLACK OR AFRICAN AM.** – Provide details below.

African American       Jamaican       Haitian  
 Nigerian       Ethiopian       Somali

Print, for example, Ghanaian, South African, Barbadian, etc. ⌘

**ASIAN** – Provide details below.

Chinese       Filipino       Asian Indian  
 Vietnamese       Korean       Japanese

Print, for example, Pakistani, Cambodian, Hmong, etc. ⌘

**AMERICAN INDIAN OR ALASKA NATIVE** – Provide details below.

American Indian       Alaska Native       Central or South American Indian

Print, for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat, Nome Eskimo Community, etc. ⌘

**MIDDLE EASTERN OR NORTH AFRICAN** – Provide details below.

Lebanese       Iranian       Egyptian  
 Syrian       Moroccan       Algerian

Print, for example, Israeli, Iraqi, Tunisian, etc. ⌘

**NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER** – Provide details below.

Native Hawaiian       Samoan       Chamorro  
 Tongan       Fijian       Marshallese

Print, for example, Palauan, Tahitian, Chuukese, etc. ⌘

**SOME OTHER RACE, ETHNICITY, OR ORIGIN** – Print below. ⌘

NALEO Educational Fund agrees with the Bureau's careful assessment. First, the Census Bureau's extensive testing has repeatedly achieved lower nonresponse rates to combined than to separate questions. "Some other race" responses also decline dramatically when a combined question is substituted for separate race and ethnicity questions. Whereas the Bureau conducted the AQE solely on paper, it tested responses received both on paper and through the internet in the NCT. In addition, while the AQE tested questions only in English, the NCT tested them in both English and Spanish. This testing confirmed that a combined question resulted in lower nonresponse rates regardless of the medium used to answer the questionnaire, or whether respondents used English- or Spanish-language questionnaires.

The Census Bureau also observed that combining race and ethnicity questions in the AQE and NCT led fewer Latinos to choose White as their race, and instead resulted in majorities of Latinos self-identifying solely in the Latino/Hispanic category and not with any of the other specific racial groups in the existing Census question. Latino NCT respondents were more likely overall to convey their Latino ethnicity in response to a combined question than if they received a questionnaire with separate race and ethnicity questions.

The Bureau also examined the consistency between respondents' answers to various question formats and the ways in which they self-identify through focused interviews with individuals. These interviews also supported the Bureau's contention that combined question formats are a better approach than separate questions to eliciting how Latinos define themselves.

Finally, when checkboxes and an optional write-in area immediately follow broad race and ethnicity categories, a combined question is as effective as separate questions in prompting survey respondents to provide detailed information about their national origins and ethnicities. This result holds true for most racial and ethnic classifications, except with respect to the Bureau's attempted use of three very broad geographic checkboxes to elicit detailed reporting from American Indian and Alaska Native respondents on paper questionnaires. The Bureau must revisit alternative format approaches to address this issue.

Results of the AQE and NCT indicate that the adoption of a combined question format will not result in the loss of any necessary data that would have been collected with separate questions. For example, the Census Bureau did not find any statistically significant differences in the rates at which respondents indicated Afro-Latino identity whether they were responding to separate or combined questions. Similarly, the percentage of people who self-identified as both Latino and White in a combined question – about 15% of all Latinos – was consistent with the percentage of Latinos who affirmed their White identity in a post-survey interview associated with questionnaires that employed separate questions. As noted, combined questions produce no decline, and instead a slight increase, in the percentages of survey respondents identifying as Latino. In sum, Latinos provide the most accurate reports of their self-identification in response to a combined question, whether or not they identify with one or more of the specific categories set forth in the Census race question.

In addition, the combined question design strategy proposed by the Bureau shows promise for collecting more accurate and nuanced data on Latinos who identify with more than one national origin or sub-group. The initial instructions for the combined question direct respondents to mark all categories that apply, and respondents who indicate they are Latino can check more than one of the checkboxes for the national origin or sub-groups listed, as well as writing in other national origins where indicated. The 2010 Census question on Hispanic origin did not permit the reporting of multiple Latino national origin or sub-groups.

### **Prospective benefits of combined question format extend beyond the Latino community:**

The Bureau also found that it collected the most detailed and complete responses from people of all races and ethnicities when it queried them using a single question with detailed checkboxes. Secondary interviews conducted after circulation of the questionnaire revealed that, across race and ethnicity categories, the identities described by respondents were generally more consistent with their answers to combined rather than to separate questions about race and ethnicity. In this connection, it should also be noted that the Bureau's combined question format includes a new "Middle Eastern or North African" category to improve the accuracy of data collected about population groups from these origins.

The Census Bureau's tests showed that in response to combined questions, larger percentages of respondents reported identifying with multiple racial and ethnic groups. In the NCT, for example, there were "similar or higher percentages of multiple-group reporting within the combined question format for Black, Asian, American Indian/Alaska Native, and Native Hawaiian/Pacific Islander groups, compared with the separate questions approach." The Bureau believes this may be because when survey respondents see all of the race and ethnicity categories grouped together in one question, they more clearly understand that they can and should choose all of the categories that apply to them.

## **Outstanding Questions for the Census Bureau and OMB**

While NALEO Educational Fund is confident that a single question about race and ethnicity can obtain optimally detailed and accurate responses from Latinos, we recognize that outstanding issues remain to be addressed by the Census Bureau and other government agencies responsible for federal data collection. Before the Census Bureau adopts a combined question format for the decennial questionnaire and other surveys, it must resolve these concerns.

First, the Census Bureau must determine how it will tabulate and report data based on responses to a combined question, and share its plans with stakeholders. For example, use of a combined question may increase the percentage of survey respondents who accurately report belonging to more than one major race or ethnicity category. Under some circumstances, the Census Bureau reports these data in one single "two or more races" category. However, if the combined question approach increases the diversity of responses by individuals who identify with more than one race or ethnicity, then presenting data on those individuals in a single category may not be an optimal approach.

Moreover, as the Bureau collects multiple responses indicating both membership in broad race and ethnicity categories and membership in smaller national origin, ethnic, and tribal groups, the Bureau will possess information about individuals' backgrounds that may be very complex and include many details. The Census Bureau must decide, and clearly communicate, the level of detail it can feasibly make available in various products and publications, striving to provide as much precise data as possible. To the extent possible consistent with sound statistical practices, the Bureau should also disaggregate these data by the different individual national origin and other ethnic groups.

If the Bureau adopts the proposed combined question approach, it could also continue to face issues regarding the comparability of its data with the OMB standards. The Bureau's proposed combined question continues to allow respondents to indicate "Some other race, ethnicity, or origin," as one of their choices. Thus, without a corresponding OMB category, the Bureau will continue to need to assign a racial or ethnic category to respondents who check this box. Based on the Bureau's NCT research, the number of individuals who choose this category should be significantly smaller than those who indicated being "Some other race" under the two separate question approach.

The Bureau should closely examine the characteristics of respondents who chose the “Some other race, ethnicity, or origin,” option under the combined question approach in the NCT to guide its work on this issue.

In addition, the Bureau must develop plans to ensure the comparability of data to historical statistics. With two separate questions on Hispanic origin and race, the Bureau has been able to present data on Latinos which also categorizes Latinos by race - either the specific racial category selected by Latino respondents, or the category the Bureau assigned the Latino respondent if he or she did not choose one of those specific categories.

However, with the combined question approach, there is likely to be a significant number of respondents who indicate that they self-identify solely as Latino. Thus, there would be no presentation of other specific “racial” categories for these respondents. While this approach may more closely reflect the self-identification of Latinos, it creates challenges for the comparability of data collected under the combined question approach and the separate questions approach. Thus, the Bureau may need to plan for and allocate resources to adapt old or new datasets for comparability, or consider developing tools that enable data users to convert old and new datasets into a common format.

## Emerging Issues and Next Steps

The Census Bureau faces some important milestones and emerging issues as it moves forward with potential redesign of the questions on Hispanic origin and race. First, the OMB is also considering revisions to its standards relating to race and ethnicity in federal data. The OMB established a Federal Interagency Working Group to address this potential revision, and in March 2017, the Working Group published a report and proposals on its standards, which will be finalized by December 31, 2017. NALEO Educational Fund provided comments to the Working Group on its proposals, and will continue to advocate for standards that promote sound and accurate federal data collection about Latinos and other population groups. In addition, by March 31, 2018, the Bureau must submit to Congress the final wording of the questions for Census 2020, so it must finish any remaining evaluation or assessment well before that date.

### **The critical need for robust Census 2020 Outreach:**

Finally, it is critical that the Census Bureau conduct and help coordinate robust outreach during Census 2020 to ensure that Latinos understand what is likely to be a new design for the Hispanic origin and race question. This outreach should emphasize that respondents should answer the question in the manner they feel best reflects their self-identification. It should also highlight the fact that respondents can choose to mark more than one racial or ethnic category, and can indicate more than one Latino national origin or subgroup. The re-design of the Census questions on Hispanic origin and race will only produce more accurate data if Latinos participate in Census 2020 and fill out their questionnaires in an informed manner.

For more information about the NALEO Educational Fund’s  
Census policy development efforts, please contact

**Ms. Rosalind Gold**, [rgold@naleo.org](mailto:rgold@naleo.org) or (213) 747-7606; or  
**Ms. Erin Hustings**, [ehustings@naleo.org](mailto:ehustings@naleo.org) or (202) 546-2536.