

# Opinion: Census questions bring about different responses

By Jennifer Lee

Like most Americans, I spent most of my life not appreciating the herculean effort the U.S. Census Bureau undertakes every 10 years.

Since its inception in 1790, the U.S. Census has aimed to count every living person in the country, and the stakes are high. The results of the census determine the allocation of hundreds of billions of federal dollars, which affect every slice of American life.

In order to do so, the Census must ask Americans the right questions—and give them the right options for their answers. It seems relatively simple, but—as I learned in 2013, when I became a member of the Committee on Population Statistics of the Population Association of America—the undertaking is so enormous that the planning for the 2020 Census began even before the completion of the 2010 Census. In 2010, the Census Bureau launched the Alternative Questionnaire Experiment (AQE) to compare different Census questionnaire design strategies. Five years later came the National Content Test (NCT), in which different questionnaires were sent to a statistically representative sample of approximately 1.2 million households in the United States and Puerto Rico.

I had the opportunity to review the results of both tests and assess which questionnaire design results in the most accurate count of the U.S. population. That meant taking three interrelated components into consideration. The first is increased reporting: Which questions were people most likely to answer? The second is decreased non-reporting: Which

questions were more likely to get groups who are susceptible to non-reporting (including poor families who get evicted, immigrants who do not read or understand English, and undocumented migrants who may fear government officials) to respond? The third is increased, detailed reporting: Which questions yield more information about the respondents?

The design of a question itself affects how people answer it. Take the race and ethnicity question. People who identify as Asian or Hispanic answer it differently depending on how it is presented on the Census form.

In the 2010 Census, Hispanic origin and race were listed as two separate questions. In both the AQE and NTC, the Census Bureau tested the option of combining race and Hispanic origin into one question, which they refer to as the “combined format.” In addition, they tested which combined format would elicit the most detailed reporting on origin.

One option was to list the racial categories only, with an option to write in their detailed origin. A second option was to list racial categories and also provide check boxes denoting examples of detailed origin, along with the option to write in one’s origin.

Figure 1 – Census Race 2010

Figure 2 – Census Race Combined With Write In Only

Figure 3 – Census Race Combined with Detailed Check Boxes

More than 70 percent of self-identified Hispanics said they were Hispanic when they were offered Hispanic as a race option (the combined option). When they are not presented with this option, as in the 2010 Census, self-identified Hispanics are more likely to check “some other race” or mark two or more races. In short, the combined option—in which Hispanic is listed as a race category—more easily allows Hispanics to accurately report their Hispanic identity. Moreover, when

Hispanics are offered the combined option, they are significantly less likely to mark "some other race" or two or more races to self-identify. Both results indicate more accurate reporting on the part of Hispanics.

Moreover, Asians were most likely to mark their race, including their detailed race, when they are provided with a check box to mark their national origin (for example, Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese). When these check boxes are removed, however, and Asians are presented with only a space to write in their national origin, they are less likely to report it. The difference is significant. The check-box format yielded a 97.4 percent response rate among Asian-Americans, and plummeted to 92.6 percent when they were provided only with a write-in option.

Detailed reporting among Asians is critical because it allows researchers to disaggregate data, which is essential to identifying health, educational, and economic disparities among Asian ethnic groups.

Such disaggregation may sound technical and mathematical, but it can have profound human impacts. For example, having data specific to different sub-groups on disease rates, health insurance coverage rates, and birth and death rates can allow policy makers and community organizations to make more informed decisions about how to best serve these populations.

Some Asian ethnic groups are more susceptible to certain health risks: Men and women of Vietnamese origin experience the highest rates of lung cancer among all Asian American subgroups, while men and women of Korean origin have some of the highest colorectal cancer rates. Such data can guide outreach on health insurance coverage; while 13 percent of Asian Americans lack health insurance, the rate is as high as 20 percent among Koreans.

In California, there's been broad recognition of the

importance of breaking out such data. Last fall, Gov. Jerry Brown, signed legislation directing the Department of Public Health to disaggregate data for the Asian American, Native Hawaiian, and Pacific Islander populations on or after July 1, 2022. Following suit, the University of California and California State University have agreed to begin releasing disaggregated data on admissions, enrollment, and graduation rates—data that will help to unveil the wide disparity in educational attainment among Asian Americans.

Data disaggregation is a powerful weapon to dismantle the dominant narrative of Asian Americans as the model minority, which has resulted in their exclusion from policy debates on poverty, health care, and education. While Asian Americans may be touted as academic high achievers, one-third of Cambodians, Laotians, and Hmong do not graduate from high school. Data disaggregation exposes these gaping differences among Asian ethnic groups, and points to the dire need for the federal resources to help boost the educational outcomes of these groups, which are essential to immigrant and second-generation integration.

If the 2020 Census provides only a write-in option to list one's origin, we will lose a lot of disaggregated data, and be unable to identify the stark differences among U.S. Asians. We will also miss a great deal of information on the country's growing and increasingly diverse Hispanic population.

*Jennifer Lee is a chancellor's fellow and professor of sociology at UC Irvine. In July, she will join the Department of Sociology at Columbia University.*