Reflecting Race? The Role of the Census in U.S. Racial Discourse, 1850-1960

Aliya Saperstein, *University of Oregon*

Abstract

The purpose of the U.S. Census is often described as holding up a mirror to society; the population count and associated demographics reflect the reality – and diversity – of the American population. A less rosy view suggests that the census creates the reality that it counts as much as, if not more than, it catalogues objective facts about the population. Through a content analysis of major U.S. newspapers, I examine the relationship between public discourse about race and the racial categories employed in the decennial census. Do the ethnoracial terms in the census come from common usage (measured by their appearance in news media prior to the date they appeared in the census), does the use of a given term in the census later prompt its adoption by media, or are the terms in the census largely removed from the way Americans talk about race in their everyday lives?

Kenneth Prewitt, director of the U.S. Census Bureau during Census 2000, has often described the purpose of the U.S. Census as holding up a mirror to society; the population count and associated demographics reflect the reality – and diversity – of the American population. Census data allow us to see ourselves as we are and give us tools to plan for the future, fight injustice and hold politicians accountable for their actions.

A less rosy view of the role of the census in American life suggests that the census creates the reality that it counts just as much as, if not more than, it catalogues objective facts about the population. This point has been argued especially vehemently when it comes to counting Americans by race and ethnicity (Nobles 2000; Skerry 2000; Hochschild and Powell 2008). Though racial and ethnic categories are now named and defined in consultation with advisory boards selected to represent various populations, this has not long been the case. The input of the populations to be counted was explicitly sought after 1970, in part because of the change from enumerator classification to self-identification on the census questionnaires. Prior to that point, the process through which the Census Bureau determined how many categories to use and what to call them was partly internal and partly subject to the whims of congressmen, lobbied – at least around the turn of the 20th century – by the (pseudo) scientists of the time (Anderson 1988). This resulted in such peculiarities as the census counting "quadroons and octoroons" in 1890 but not after, and counting "Hindus" in 1920, when most Americans of Indian origin were actually Sikhs (Snipp 2003).

This study seeks to examine the relationship between public discourse about race and the racial categories employed in the decennial census through a content analysis of U.S. major newspapers. Do the ethnoracial terms in the census come from common usage (as measured by their appearance in newspapers prior to the date they appeared in the census), does the use of a given term in the census later prompt its adoption by media, or are the terms in the census largely removed from the way Americans talk about race? I begin my analysis with the addition of the term "mulatto" in the 1850 census and end with the addition of "Hawaiian," "Aleut" and "Eskimo" in 1960.

Preliminary evidence, drawing on data from the *Los Angeles Times* archives, suggests support for each of these possibilities, depending on the racial term in question. As there is likely a high degree of regional variability in racial discourse, due to unique patterns of settlement, migration and identification, these results should be interpreted cautiously. Additional data collection using searchable archives from the *New York Times, the Chicago Tribune, the Miami Herald* and *the New Orleans Times-Picayune* will be conducted to determine whether the patterns in the Los Angeles data are generalizeable to other major metro areas.

Data and Methods

The *Los Angeles Times* began publishing in 1881 and its archives are searchable online through Pro Quest until 1990. In the first stage of data collection, keyword searches were conducted

using each of the 14 racial category terms that were introduced in the U.S. Census between 1850 and 1960.[1] Decade-by-decade counts were recorded for each term, including the total number of documents in which the term appeared, the number of classified advertisements and the total number of articles. Unless otherwise specified, the number of articles in which a given racial term appeared is used in the analysis below. The total number of documents, classified ads and articles was also recorded for each decade in order to adjust for increases in use that may have resulted from increases (or decreases) in the amount of published content during a particular time period. For several of the terms, including Filipino and Hindu, alternative spellings were also searched and counted separately (e.g., Pilipino or Hindoo).

Table 1. History of U.S. Census Racial Terms

1st Year in Census	Racial Term
1850	Mulatto
1860	Indian
1870	Chinese
1890	Quadroon
	Octoroon
	Japanese
1920	Filipino
	Hindu
	Korean
1930	Negro
	Mexican
1960	Hawaiian
	Aleut
	Eskimo

Source: Nobles 2000, Appendix A

In the second stage of data collection, decade-by-decade searches were conducted for the number of documents, classified ads and articles that contained the words "race," "racial," "ethnic" or "ethnicity."[2] These counts formed the basis for generating a sample of articles about racial and ethnic distinctions from which potential racial categories could be gathered that might not have been included in the census but were in popular use -- at least in the media -- during the study

---

[1] Searches were not conducted for the terms "white" and "black" alone due to the frequency with which these words are used in non-racial contexts. However, counts were compiled for the phrases "white race" and "black race." These explicit phrases were used fairly rarely; in the 1920s, the "white race" appeared in 238 articles (.03 percent of all articles in the decade) while the "black race" appeared in 51 (.007 percent).

[2] The word "race" often appears in the newspaper in generally non-racial contexts, such as reporting the results of sporting events and elections. To minimize the inclusion of these non-relevant articles, the keyword search was conducted using the maximum number of limiting terms (e.g., AND NOT) allowed by ProQuest. As this strategy would also remove articles in which the race of a competitor or candidate was mentioned, I chose to use limiting terms associated only with sports and not politics under the assumption that race in the sense in which I am interested would be more likely to occur in an election race than, for example, a horse race. Examples of such limiting terms included: pony, yacht, track, mile and Olympics. Even with these efforts, on average, about one-third of the sample of articles for each decade included articles that turned out to be non-relevant for the current purposes.

period. For each decade between 1881-9 and 1920-9, a total of 100 articles was sampled at random from this population of articles.[3] These articles were then read for content, paying particular attention to the terms used to describe individuals or groups in a racial or ethnic context. Decade-by-decade counts based on this set of "non-Census" racial terms were also conducted to assess the popularity of these alternate terms and how they varied over time during the study period.

The following excerpt from a June 6, 1905 article entitled "The Problem of Immigration" provides an example of how these articles were coded and the list of non-Census ethnoracial terms compiled.

> Japan, like China, has a dense population and it would probably be a relief to its overcrowded cities if from 5,000,000 to 10,000,000 of its surplus people should emigrate. Naturally, if unhindered, they would swarm on our western shores, and after taking the place of members of the Caucasian race in shop and field it would at length be found that they were an unwelcome element ... No one having solicitude for the future welfare of California can be indifferent to a great influx of either branch of the Mongol race.

Based on this text, both "Caucasian" and "Mongol" would be added to the list of non-Census racial terms. For an abridged version of the list, see Table A1.

Preliminary Results

Detailed analysis of the non-Census racial terms will be included in the final manuscript and presented in contrast to those terms the census elected to employ in its enumerations. Here, I present a preliminary overview of trends in the use of census racial terminology. I also briefly explore trends in the use of the general terms, such as "race" and "ethnicity."

Creating or Reflecting Reality?

Levels of popular use vary widely across the set of racial categories introduced into the census between 1850 and 1960, from zero articles mentioning the term "Filipino" in the 1880s to 20,876 articles that mentioned "Japanese" in the 1940s. Proportionally, the media's attention to these racial categories ranged from .00005 percent of all articles in the 1950s using the term "octoroon" to upwards of 3 percent of all articles mentioning "Japanese" during the 1940s (during and immediately after World War II) and "Chinese" during the 1880s and 90s (in the years immediately before and following the passing the Chinese Exclusion Act of 1882).

---

[3] A random number generator was used to select from the total number of articles in a given time period. ProQuest limited viewing full text to 2,000 articles per search, so in most cases 10 articles were sampled at random from each calendar year to create a total sample of 100 articles per decade.

Examining the set of racial census categories as a whole, there is no clear pattern as to whether terms were popular in the media prior to their inclusion in the census, whether they were only picked up after their adoption or whether the census terminology was rarely used in printed public discourse. Some terms, such as "Negro" and "Mexican," were commonly seen in print many decades before their inclusion in the census. (On average they appeared in 1.3 percent and 2.1 percent of all articles, respectively, across the nine decades.) Other terms, such as "quadroon," "octoroon," "mulatto" and "Hindu," barely registered in the pages of the *Los Angeles Times* at any time during this period (See Figure 1).

It is clear, however, that in most cases usage of a given term gets a significant "bump" in the decade in which it first appears in the census. (The exceptions are "Octoroon" and "Korean.") Sometimes this boost in popularity lingers for a few decades following the term's inaugural census appearance -- as seems to be the case for "Hindu" and "Mulatto," as well as for "Filipino" and "Chinese" (See Figure 2). Nevertheless, the largest spikes in the use of particular racial terms are more likely to correspond to other political events, particularly the many wars during this period (Spanish-American in the 1890s, Korean in the 1950s, etc.).[4]

Overall, the patterns in terminology popularity from the *Los Angeles Times* suggest that the categories used in the census between 1850 and 1960 did not reflect well-recognized groups that featured prominently in public discourse at a given point in time. For example, the pseudo-scientific debates about the viability of the offspring of mixed African and European unions that prompted the adoption of "quadroon" and "octoroon" in the 1890 census apparently did not capture the interest of newspaper editors, reporters and/or readers at all -- at least on the West Coast. Similarly, though concerns about the assimilation and integration of immigrants supposedly justified the inclusion of new racial categories around the turn of the 20th century (Anderson 1988, Nobles 2000), widespread public conversation did not seem to be occurring about Filipinos, Hindus, Koreans, or even the Japanese, until after they were given a census racial category. Meanwhile, Mexicans, ones of the categories that received the most consistent coverage in the *Los Angeles Times* during this period appeared only fleetingly in the census in 1930 and did not re-appear explicitly again until 1980 when the census bureau added a separate question on Hispanic origin.

Race or ethnicity?

In trying to establish a population of articles in which racial terminology could be found, separate decade-by-decade searches were conducted for the terms "race," "racial," "ethnic" and "ethnicity." There are several trends in the popularity of these terms in the *Los Angeles Times* that are worthy of note. While academics have debated the appropriateness of using the terms "race" or "ethnicity" to describe group boundaries in the United States, and around the world, the

---

[4] If, for example, a search for the term "Korean" is conducted with the limiting term "AND NOT war", the proportion of articles that use Korean in the 1950s drops from 13,588 to 4,869, or from 1.6 percent of articles in the decade to .6 percent.

popular vote clearly goes to using the terms "race" and "racial" rather than "ethnic" or "ethnicity" to describe difference among Americans.

First, the term "ethnicity" does not appear in a single document until the 1960s. The term "ethnic" first appears in 1886, but the number of articles that include the term hovers around 20 per decade until the 1950s when there are 117 articles in which the word "ethnic" appears. ("Ethnic" eventually hit its peak use in the 1990s with nearly 25,000 articles, or 1.6 percent of all articles printed that decade.) The term "racial" is far more popular throughout the study period. It appears in every decade and is used in more than 1,200 articles in the 1920s. Peak use of the word "racial" in articles occurs in the 1960s, when nearly 12,000 articles -- or 1 percent of all articles in the decade -- include the term at least once.

The term "race" is several orders of magnitude more popular than any of the other three terms, though the counts are tempered by the caution noted above regarding frequency of use in non-racial contexts such as sporting events and elections. Without using search restrictions, the word "race" appears in anywhere from 4.4 to 7.4 percent of all articles in the *Los Angeles Times* between 1881 and 1959. It is most prevalent in the 1930s (48,000 articles) and least prevalent in the 1880s (3,100 articles). As many as half of these articles would be considered non-relevant to the current study, but including the limiting terms noted above in these "race" searches does not change conclusions about when the term was most and least popular in the Los Angeles media.

In future research, I plan to read and code samples of articles drawn from each of these searches to determine which specific ethnoracial terms (e.g., Mexican, Negro) are used in which context. This would demonstrate which groups are seen as representing "ethnic" versus "racial" differences and whether that has changed over time.

References

ANDERSON, MARGO 1988 *The American Census: A Social History*, New Haven, CT: Yale University Press

HOCHSCHILD, JENNIFER and BREANNA MAREA POWELL 2008 'Racial reorganization and the U.S. Census 1850-1930: Mulattoes, half-breeds, mixed parentage, Hindoos and the Mexican race,' *Studies in American Political Development* vol. 22, pp. 59-96

NOBLES, MELISSA 2000 *Shades of Citizenship: Race and the Census in Modern Politics*, Stanford, CA: Stanford University Press

SKERRY PETER 2000 *Counting on the Census? Race, Group Identity and the Evasion of Politics,* Washington D.C.: The Brookings Institute

SNIPP, C. MATTHEW 2003 'Racial measurement in the American census: past practices and implications for the future', *Annual Review of Sociology* vol. 29, pp. 563-88

Table A1. Examples of "Non-Census" Ethnoracial Terminology

Jap OR Japs
Chinaman OR Chinamen
Asian-American
Asiatic
Oriental OR Orientals
Celestials
Coolie
Malay OR Malays
Mongolian OR Mongoloid OR Mongol
Kanaka
West Indian
Colored
Negroid
Afghan OR Afghans
Jewish OR Jew OR Jews OR Hebrews
Anglo-Saxon
Caucasian Or Caucasoid
Aryan or Aryans
Nordic OR Nordics
Alpine OR Alpines
Mediterraneans
Celt OR Celts OR Celtic
Germanic OR German race
Teuton OR Teutonic
Slav OR Slavic OR Slavs
Gypsies
Yellow race
Red race
Brown race
Black race
White race
White ethnic OR White ethnics
Mixed race
Mixed blood
Mestizo
Multiracial

Source: *Los Angeles Times* historical archive, 1881-1929
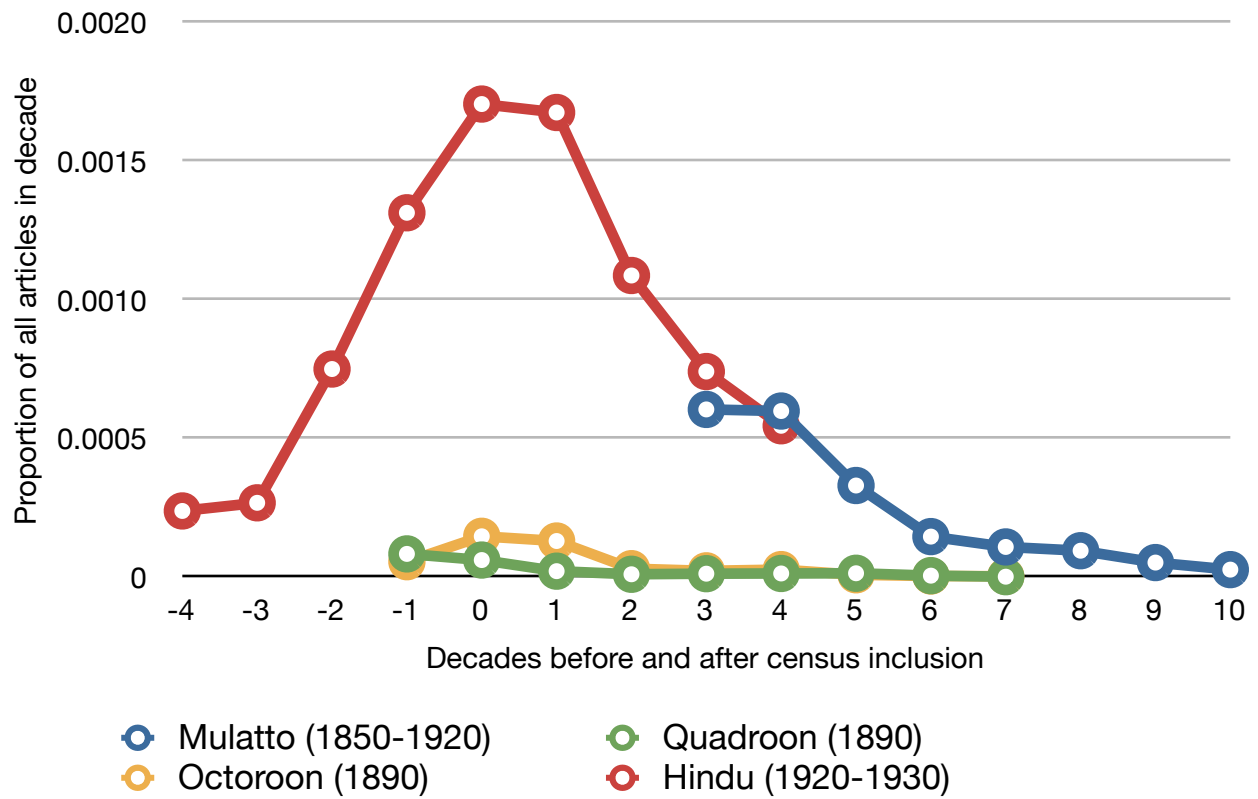
Figure 1. Usage trends in terms that came and went

Figure 2. Usage trends in terms to monitor allegedly problematic populations