

Robbin, A. (1999). The problematic status of statistics on race and ethnicity: An "imperfect representation of reality." *Journal of Government Information*, 26(5), 467-483.



## THE PROBLEMATIC STATUS OF U.S. STATISTICS ON RACE AND ETHNICITY: AN “IMPERFECT REPRESENTATION OF REALITY”

ALICE ROBBIN\*

School of Information Studies, Florida State University, 240 Louis Shores Building,  
Tallahassee, FL 43406-2100 USA. Internet: <arobbin@mailier.fsu.edu>

**Abstract**— This article extends Stratford’s brief observations about the problematic status of racial and ethnic group statistics to a discussion of the relationship among these statistics, public policy, and the conceptual status of race and ethnicity. Federal statistics are organizational products that are socially constructed. They represent the implementation of public policies that govern political, social, and economic life. It is the interaction between politics and the subjective meaning of race and ethnicity that is responsible for the continual modification of racial and ethnic group statistics. The article discusses the premises on which racial and ethnic group statistics have been based and illustrates how they were implemented in the instructions of the decennial censuses for classifying the race and ethnicity of the population. The article then summarizes some of the empirical evidence from recent research conducted by federal agencies and social scientists to show that racial and ethnic group statistics produced by government record keeping systems have no objective status. The meaning of race and ethnicity is contextual, situational, and subjective, and, thus, how respondents and observers define these concepts has significant consequences for the quality of federal statistics. © 1999 Elsevier Science Ltd. All rights reserved.

**Keywords**— Federal statistics, Decennial census, Race, Ethnicity, Minority groups

### INTRODUCTION

The routine production of statistical information reinforces a sense that the measures are real, the properties of categories invariant, and their meaning unproblematic. The contrary is, however, the reality—as Stratford asserts in a recent article about the difficulties that the information specialist faces in responding to reference queries about government statistics [1]. Their “problematic status,” in particular, a lack of consistency in the definitions and methodology, creates significant difficulties for interpretation, analysis, and their application in public policy.

This article extends Stratford’s observations about statistics on race and ethnicity to a discussion of the relationship among these statistics, public policy, and the conceptual status of race and ethnicity. It is the interaction between politics and the subjective meaning of race and ethnicity that is responsible for the continual modification of the racial and ethnic group statistics, and explains why, not unexpectedly, the federal classification of data on race and ethnicity has been the subject of political debate and controversy.

---

\*Alice Robbin is an assistant professor in the School of Information Studies at Florida State University, where she teaches courses in government documents, information policy, research methods, and the societal implications of the information age.

The thesis of this article is that these statistics are socially constructed organizational products. Part one discusses some of the premises on which statistics on race and ethnicity have been based and their relationship to public policy. The technology of census enumeration—methodology and classification—reproduces historical and current public policy and legitimates state action for future decisions. How these premises were implemented as public policies that governed social, political, and economic life and were subsequently incorporated as instructions to enumerators of the decennial censuses for classifying the race and ethnicity of the population is illustrated. Race and ethnicity items on the decennial census schedules are deconstructed, and inferences about meaning are drawn by situating the text in its larger socio-political context. Part two reinforces Stratford's observations about the problematic status of these statistics. Empirical evidence from recent research conducted by federal agencies and social scientists is analyzed to show that statistics on race and ethnicity produced by government record keeping systems have no objective status. The meaning of race and ethnicity is socially constructed, contextual, situational, and subjective. Thus, how these concepts are defined by respondents and observers have significant consequences for the quality of the measurements in decennial census enumerations, surveys, and administrative record keeping systems maintained by the federal government [2].

### **POLITICAL BASES OF RACIAL AND ETHNIC GROUP STATISTICS**

Policies related to property rights, political representation, and citizenship in the polity have always motivated administrative practices for collecting data on race and ethnicity [3]. Each policy domain has reinforced the other. Census enumeration has been the government's instrument for determining a group's status in the polity and society. By defining who is counted and how people are classified, the census has provided the statistical information on which public policy is designed, programs implemented, and federal, state, and local statutes enforced [4].

Racial classification in the United States has been premised on a theory that links skin color to social status, and on race as an invariant property of an individual. "Perceived behavioral features and differences in intellect served as the basis for the ranking, in terms of superiority, of races," and "early natural history approaches to racial classification supported these rankings and the implications for behavior" [5]. Government-established racial categories reflected efforts to distinguish who was not of northern European ancestry and could be classified as "nonwhite" and, thus, inferior (i.e., the whiter the skin, the higher the status) [6]. No amount of white ancestry, except 100 percent, could classify a person as "White." The "one drop of blood" (hypodescent) rule has governed, either explicitly or implicitly, the classification of race for all "nonwhite" persons, naming conventions, and observer perception of the proportion of African blood. This rule has perhaps been the single most important social organizing concept in the United States, privileging one class ("White") over another, whether the assigned label was "Colored," "Not-White," "Other," "Mulatto," or "Mestizo." Efforts to distinguish precisely the amount of Negroid blood in all races reinforced public policies that assigned political rights on the basis of race as well as the practice of exclusionary politics [7].

### **EARLY RACIAL AND ETHNIC CLASSIFICATION**

The United States has classified race since its first census in 1790 [8]. Ethnicity has been classified since the 1850 census with a series of questions that have asked citizens

about their origins. Definitions and naming conventions applied to social groups have functioned to reinforce law, policy, and practices. Politics and social and economic conditions have contributed to the appearance and disappearance of racial and ethnic group categories [9]. Table 1 shows the population items related to race and ethnicity in the decennial censuses between 1790 and 1990.

The earliest censuses distinguished majority and minority populations by property for purposes of taxation and political representation, and naming conventions denoted civil or legal status. For example, “Indians” (classified as “All Others” in the early censuses) were taxed, but had no representation. People of African ancestry were either classified as “slaves” (property) and had no political representation, or as “coloured” if freemen and counted as three-fifths of a White person.

Immigration and naturalization laws contained explicit language that identified particular ethnic groups in racial terms and as ineligible for citizenship. As early as the 1820s, the formal definition of former residents of Latin America was fixed by national citizenship (e.g., inhabitant of Mexico is “Mexican”), but the passage of the Foreign Miner’s Tax in 1850 excluded both foreign and native born Latin American miners from claiming rights. Subsequently, all intra-Latin American differences were ignored, and both native and foreign born Latin Americans became “foreigners” [10]. Motivated by sinophobia, nativist sentiment, and dislocations in the economy, the additions of a “Chinese” category in the 1870 census and a “Japanese” category in the 1890 census were designed for exclusionary purposes.

## CATEGORY CLASSIFICATION AND SOCIAL INTEGRATION

Category classification has also been employed to legitimate a political relationship that social groups have historically held with the federal government. Both Native Hawaiians and American Indians, for example, have been officially classified as a “race” by the U.S. government even though both groups do not conceive of themselves as a “race” [11]. Instead, they have argued, such classification affirms a political relationship with the U.S. government through treaty compacts over centuries of conflict and disputes.

The evolution of category naming conventions from “non-white” to the “White” race or an ethnic group indicated eligibility for membership in the polity and the status of its social integration. This is illustrated by changes from racial to ethnic category assignment during the nineteenth and twentieth centuries for nearly all non-northern European whites and non-whites [12]. For example, at one time, Hindus, Indians, Filipinos, and other “Orientals” were treated as members of a “non-white” race. The Irish became “White” [13], as did Italians. Syrians were classified as “Asiatic” during the nineteenth century until they became “White” in the early twentieth century; however, even as late as 1970, instructions for completing the census informed the enumerator that “Moslem” or “Brown” was to be changed to “White” [14]. Mexicans were first classified as “White” following the 1848 Treaty of Guadalupe Hidalgo. They were explicitly enumerated in the 1930 decennial census as a separate “race,” and evolved in 1940 to an ethnic category. In the 1960 and 1970 censuses they were classified as “persons of Spanish mother tongue” and “White persons of Spanish surname” [15]. In 1960, Puerto Ricans, Mexicans, or other persons of Latin descent were to be classified as “White” unless they were “definitely Negro, Indian, or some other race.” See Table 2 for the evolution of Spanish-speaking naming conventions and ethnic classification in the census schedules between 1930 and 1990.

Table 1. Population items related to race and ethnicity on general schedules of decennial censuses: 1790–1990

Items	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990
Race	X <sup>1</sup>	X	X	X	X	X	X	X	X	X	X	X	X	X							
Hispanic origin																					
Citizenship <sup>2</sup>									X		X	X	X	X	X	X	X	X	Xs	Xs	X
Year of entry <sup>3</sup>				X <sup>1</sup>	X <sup>1</sup>						X	X	X	X	X	X	X	X	Xs	Xs	Xs
Ancestry																					Xs
Place of birth <sup>4</sup>							X <sup>1</sup>	X <sup>1</sup>	X	X	X	X	X	X	X	Xs	Xs	Xs	Xs	Xs	Xs
Place of birth of parents <sup>5</sup>									X	X	X	X	X	X	X	Xs	Xs	Xs	Xs	Xs	Xs
Language: Non-English language <sup>6</sup>									X		X	X	X	X	X	X	X	X	X	Xs	Xs
Mother tongue <sup>7</sup>											X	X	X	X	X	X	Xs	Xs	Xs	Xs	Xs

X = Item asked of entire population; Xs = Item asked of sample.  
 Sources: Adapted from Jorge del Pinal and Susan J. Lapham, "Impact of Ethnic Data Needs in the United States," in *Challenges of Measuring an Ethnic World* (Proceedings of the Joint Canada-United States Conference on the Measurement of Ethnicity, April 1–3, 1992) (Washington, DC: Department of Commerce, Bureau of the Census and Ottawa: Statistics Canada, 1992), 466; U.S. Department of Commerce, Bureau of the Census, *Race and Ethnic Origin 1990* [Census of Population and Housing Content Determination Reports] (1990 CDR-6)(Washington, DC, 1991); U.S. Department of Commerce, Bureau of the Census, *Birthplace, Citizenship, Year of Entry, and Language* [1990 Census of Population and Housing Content Determination Reports](1990 CDR-7)(Washington, DC, 1990).

<sup>1</sup>Item asked of "Free white inhabitants" only in 1790, 1800, and 1810; "All free white and colored (Black and Mulatto) inhabitants" from 1820 to 1860.  
<sup>2</sup>In 1870, citizenship was asked only of males 21 years and older. From 1890 through 1960, except the 1960 census, citizen questions for foreign-born persons were included. In 1960, the citizenship questions were asked only of persons living in New York and Puerto Rico. In 1990, the item on citizenship was asked of all native and foreign-born persons.

<sup>3</sup>1890: "Number of years in the United States;" for foreign-born persons; 1900–1930: "Year of immigration;" of foreign born; 1970–1990: "When did...come to the United States to stay?"

<sup>4</sup>In 1960, "Place of birth" was asked on a 25 percent sample basis, but in New York and Puerto Rico on a 100-percent basis.  
<sup>5</sup>Question was only whether parents were foreign-born in 1870.

<sup>6</sup>In 1890 and 1910, question was simply whether persons could speak English. In 1910, 1980, and 1990 the question obtained information on ability to speak English. Beginning in 1880 through 1940, the supplemental Indian Division Population Schedule identified "Other than native language spoken."

<sup>7</sup>In 1910, 1920, and 1940, obtained mother tongue of respondent; in 1910 and 1920, obtained mother tongue of his/her foreign-born parents.

Table 2. Nouns related to the Spanish-speaking population on the general schedules of decennial censuses, 1930–1990<sup>1</sup>

Ethnicity/Nationality	1930	1970	1980	1990
Spanish/Hispanic		X	X	
Mexican <sup>2</sup>	X	X	X	X
Mexican-American			X	X
Chicano		X	X	
Puerto Rican		X	X	X
Cuban		X	X	X
Central or South American		X		
Other Spanish		X		
Other Spanish/Hispanic			X	
Spaniard			X	
Argentinian				X
Colombian				X
Dominican				X
Ecuadoran				X
Honduran			X	X
Nicaraguan				X
Salvadoran				X
Venezuelan				X

*Source:* Department of Commerce, Bureau of the Census, *200 Years of U.S. Census Taking: Population and Housing Questions, 1790–1990* (Washington, DC, 1989).

<sup>1</sup>Note there were no nouns describing the Spanish-speaking population on the 1940, 1950, and 1960 census schedules. The named ethnic groups appear on the general schedules in items on “Race,” “Hispanic Origin,” “Birthplace of self or parents,” or “Ancestry.”

<sup>2</sup>See endnote 10 for more explanation of the evolution of naming conventions.

## NONWHITE NAMING CONVENTIONS

Naming conventions applied to the non-white population and instructions for enumerating the population in decennial censuses and vital statistics regulations show the extent to which the federal government went to calculate the amount of non-white blood, the government’s dichotomous conception of race, and the relationship of race to political rights. This obsession with distinctions of blackness is illustrated by some of the census instructions for enumerating blacks and American Indians during the nineteenth and early part of the twentieth centuries. See Table 3 for the nouns applied to the “White” and “non-white” African American populations on the schedules of the decennial censuses for the period 1790 through 1990.

Beginning in 1850, efforts were made to distinguish the “free coloured” population, which was further differentiated into “Black” and “Mulatto.” Marshals were similarly instructed to distinguish slaves by the amount of white blood. In 1870, enumerators were instructed to, “Be particularly careful in reporting the class *Mulatto*. The word is here generic, and includes quadroons, octoroons, and all persons having any perceptible trace of African blood. Important scientific results depend upon the correct determination of this class” [16]. In 1890, enumerators were instructed to “distinguish carefully between blacks, mulattos, quadroons, and octoroons. The word ‘black’ should be used to describe those persons who have three-fourths or more black blood; ‘mulatto,’ those persons who have from three-eighths to five-eighths black blood; ‘quadroon,’

Table 3. Naming conventions applied to white and African American-origin populations on the general schedules of decennial censuses: 1790–1990<sup>1</sup>

Racial Category	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	
White	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X <sup>6</sup>	X	X	X <sup>7</sup>	X	X	X	X
Coloured	X	X	X	X	X	X																
Slave	X	X	X	X	X	X	X															
Black (Negro)							X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Mulatto							X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X <sup>8</sup>
Quadroon							X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Octoroon							X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Afro-American											X										X	X

Source: Department of Commerce, Bureau of the Census (1989).

<sup>1</sup>Terms (nouns) and spelling employed in the schedules of the U.S. Censuses of Population. The nouns appear on the general schedule in the “Race” item.

those persons who have one-fourth black blood; and 'octoroon,' those persons who have one-eighth or any trace of black blood" [17]. In 1900, "Black ('B') is defined as negro or of negro descent" [18]. In 1910, the term "black" includes all persons who are evidently full-blooded negroes, while the term 'mulatto' (Mu) includes all other persons having some proportion or perceptible trace of negro blood" [19]. In 1930, a "person of mixed White and Negro blood was to be returned as Negro, no matter how small the percentage of Negro blood" [20]. The 1930 census also instructed enumerators that, "Any mixture of White and some other race was to be reported according to the race of the parent who was not White; mixtures of colored races were to be listed according to the father's race, except Negro-Indian" [21].

Starting with the 1880 census, racial distinctions were also extended to the American Indian. Enumerators were instructed to identify whether the Indian was full- or mixed-blood ("White," "Black," or "Mulatto"). In 1900, more detailed instructions required that the enumerator define "how much" white blood an Indian had "(0, 1/2, 1/4, or 1/8) whichever fraction is nearest the truth" [22]. In 1910, in an effort to further define the amount of blood, the amount "3/4" was added, and the enumerator was instructed to consult an older member of the tribe "to verify" the information because "an Indian is sometimes of mixed blood without knowing it" [23]. In 1920, however, there was no Indian Schedule, so detailed information on blood was not obtained. In 1930, a "person of mixed White and Indian blood was to be returned as an Indian, except where the percentage of Indian blood was very small or where he or she was regarded as White in the community. For persons reported as American Indian, enumerators were instructed to indicate the degree of Indian blood." Moreover, someone who was "part Indian and part Negro also was to be listed as Negro unless the Indian blood predominated and the person was generally accepted as an Indian in the community" [24].

### **"OTHER" RACIAL IDENTITY**

Through 1960, enumerators were instructed to complete "Color" in "Race" by "observation." The 1970 Census represented the first where all answers were given by self-identification, except that the enumerator was permitted to fill in blanks by observation. The instructions for enumeration of race included a long list of possible written-in entries and how they were to be classified. "For example, 'Chicano,' 'LaRaza,' 'Mexican American,' 'Moslem,' or 'Brown' were to be changed to White, while 'Brown (Negro)' or Black would be considered as Negro or Black for census purposes" [25]. In 1980, enumerators were no longer permitted to enter race by observation, but "were instructed to ask and mark the race with which the person most closely identified. If a single response was not possible, as in the case of a racial mixture, the mother's race was to be reported. If this was not satisfactory, the first racial group given was to be entered" [26].

The "Other" as a category of racial identity has been an essential means of distinguishing the "non-white" and "white" populations [27]. In the censuses of 1790 through 1810, "all other free people" were enumerated. In 1910, instructions stated that, "for all persons not falling within one of these classes, write 'Ot' (for other), and write on the left-hand margin of the schedule the race of the person so indicated." In 1960, enumerators were instructed to classify Asian Indians as "Other" and write in "Hindu" [28]. In 1980, "Other" could include "Brown," "Mexicano," etc., unless one of the other categories was selected [29]. Use of the "Other" category has been applied through the 1990 census to classify any person who self-identified as belonging to more than one race.

In the last decades of the twentieth century color remains an important part of the national political discourse, an explicit attribute of racial identity assigned by the state and a reflection of the status of race relations in the society. As late as 1970, the Louisiana state legislature labeled anyone with 1/32 “Negro blood” as “black” [30]. Although repealed in 1983 by legislation that permitted a person to change birth records by presenting “a preponderance of evidence” to prove that the record was wrong, subsequent appeals by the applicant Guillory Phipps to modify the racial identity of her parents from “colored” to “white” in the birth certificates, were rejected in 1986—even as the Supreme Court of Louisiana granted that “individual racial designations [were] purely social and cultural perceptions” [31]. The court decision has not been overturned.

### **SKIN COLOR AND PUBLIC POLICY**

During the last two decades, internal U.S. federal agency and public debates about federal standards for classifying and reporting racial and ethnic data have explicitly acknowledged the relationship between skin color and public policy, in particular, the classification of the “non-white” populations of dark-skinned East Indian Asians, peoples of Middle Eastern descent, South and Central Americans of American Indian descent, and the increasing numbers of multiracial persons [32]. The debates that took place in the 1970s and 1990s about Statistical Policy Directive 15, the Office of Management and Budget standard for classifying data on race and ethnicity, reflected, however, major changes in federal social policy, whose origins were the legislative initiatives of the Great Society decade of the 1960s. The principles that supported a classification system for racial and ethnic data emphasized inclusion rather than exclusion in the body politic, derived from a recognition of large-scale demographic changes in the population [33], and were also attempts to rectify centuries of injustice and discriminatory practices against marginalized social groups.

Politics also influenced congressional and administrative policy decisions during this period. Great Society initiatives fostered and reinforced the rise of identity politics and the entrance of minority population interest groups into the political arena. African-Americans entered during the 1960s; Latinos in the early 1970s; and Asians and Pacific Islanders in the late 1970s. Their mobilization contributed to altering the political, economic, and social landscape [34]. Successful mobilization required, nonetheless, significant congressional support and public acknowledgment of the deprivations suffered by minority groups to prod government agencies to develop mechanisms that registered the health, political, educational, and economic status of marginalized groups [35]. Political mobilization by minority populations created “voice” in the development of administrative record keeping systems of the government that recorded information on race and ethnicity, and was directly responsible for the expansion of racial and ethnic group categories [36].

### **PROBLEMATIC STATUS OF RACIAL AND ETHNIC IDENTITY**

When the revised standards for classifying racial and ethnic data were proposed in mid-1997, they met immediate opposition from the American Anthropological Association. The Association argued that the concepts of race and ethnicity had no scientific or anthropological basis and were culturally and socially defined. The Association’s statement also alluded to a body of empirical evidence, which indicated a variety of “conceptual confusions” or “misunderstandings” about the concepts of race and ethnic-

ity. The Association also argued that the meaning of race and ethnicity was contextual, situational, and subjective, and contributed to the problematic status of the statistics. Fundamentally, two issues determine the quality of federal statistics on race and ethnicity: how both respondents and observers interpret the concepts, and how the record keepers word and sequence the questions in the instrument.

The U.S. Bureau of the Census and other investigators' more than two decades of research have supported findings about the "conceptual confusions" related to the terms "race," "ethnicity," and "ancestry." Government agencies had, however, assigned the meaning of "misreporting" to respondents' inability to classify correctly themselves into the racial and ethnic group categories. It required substantial empirical evidence to persuade government statisticians of the underlying cognitive assumptions that respondents made about the concepts of race and ethnicity.

### **CURRENT POPULATION SUPPLEMENT ON RACE AND ETHNICITY**

One of the sources of evidence is a series of 83 in-depth cognitive interviews that were conducted by the U.S. Bureau of Labor Statistics and U.S. Bureau of the Census in November 1994 [37]. These interviews were designed to test whether conceptual questions about race, ethnicity, national origin, and ancestry could be included in the 1995 Current Population Supplement (CPS) on Race and Ethnicity [38]. All three phases of the cognitive research showed that respondents experienced a conceptual confusion, no matter whether questions were revised or reordered—even among those with some years of college education. Most of the respondents interpreted "race," "ethnicity," "ancestry," and "national origin" as the same terms—a single concept, and found these terms redundant and the questions associated with them too difficult and too abstract. The conclusion reached by the McKay and de la Puente research team was that these terms were interpreted as semantically identical [39].

The research team found that very few respondents knew what the term "ethnicity" meant. Some of the respondents believed that these questions were a test of their intelligence, and some individuals thought the researchers were probing to find out something about the "ethical character of the races." The researchers also found that the word "characteristic" was interpreted to mean one's personal character. When asked to identify the specific racial group to which they belonged, some respondents replied "Christian," "Masons," "Black Muslims," or "Rebellious teenager." Non-Hispanic Whites responded to "Are you Hispanic, Latino, of Spanish origin, or none of these?" by saying, "I'm White" or "I'm American" [40].

The researchers found that some respondents were confused by a question that asked them to name an ethnic group within their racial group [41]. Respondents wondered if the researchers were asking about membership in "all-White" or "all-Black" social groups. Some respondents thought that they were being asked about being multi-racial, and one "Black" respondent asked if "naming an ethnic group within their racial group" meant distinctions about color or blackness. Even more precise questions about race, ethnicity, and ancestry, which the researchers believed would provide an anchor for the respondent's answers, demonstrated conceptual confusion and were found to be too abstract [42]. The research team finally concluded that this series of questions was too difficult for respondents, and did not include them in the national survey that was eventually fielded. Even so, inconsistencies in responses were observed during the subsequent fielding of the final version of the survey. Once again, respondents found it difficult to distinguish race, ethnicity/ethnic origin, and ancestry. Respondents of both

White and American Indian ancestry appeared to be inconsistent in identifying with one or the other race. White ethnics of multiple European ethnic origin/ancestry (e.g., both “Italian” and “Polish”) reported themselves as “multiracial” [43].

### MULTIRACIAL POPULATION

Additional evidence about the subjective meaning of the concepts was also found in a laboratory study of multiracial and multiethnic women conducted by researchers at the University of Illinois in 1996 [44]. Their study revealed that the assignment of race or ethnic group presented conceptual difficulties for multiracial persons. Subjects were inconsistent in reporting their race or ethnic group, and nearly half reported that they had self-identified in the past differently than they did now. Racial and ethnic identities were far more complex for the multiracial population. The researchers found that answers that the respondents gave were highly sensitive to the terminology used to describe a person’s race or ethnicity. Their results gave further support to the McKay and de la Puente research, which had revealed that “measurement of racial and ethnic identity [was] largely dependent on the shared use of terminology,” and that there was “considerable variability in the racial and ethnic labels selected by individuals within each subgroup” [45]. The findings of the University of Illinois research team also reinforced a growing body of empirical evidence that the concepts of race and ethnicity were culturally defined and situational.

Empirical research carried out over more than two decades has consistently shown that racial or ethnic identity is influenced by culture, context, and subjective interpretation, including the views of others. For example, Hispanic or Latino racial identity depends on language, social class, neighborhood socialization, phenotype, and phenotypic variation within families [46]. Outside the United States, particularly in Latin America, race is perceived as a broad spectrum of colors [47]. People may identify themselves as racially different, even within the same family. Only upon coming to the United States must Hispanics choose to be either black or white. Seeing neither of them as appropriate leads them to choose the residual category of “Other,” reject the OMB categories for classifying race, or self-identify by national origin [48].

There is substantial evidence that ethnic identity is fluid, flexible, and varies over the life course, as well as subjective, contextual, and situational [49]. For example, about 35 percent of the people who were interviewed in both the March 1971 and March 1972 Current Population Surveys demonstrated variation in their assigned ethnicity [50]. It is important to note that this variation also occurred because a family member could answer for the respondent (acting as a proxy) in one or the other survey—reinforcing the scientific insight about subjectivity with regard to ethnic and racial identity. Johnson hypothesized, and Alba and Chamlin and Waters independently established with their research on “White” ethnics, that reporting changes could be due to a lack of information or uncertainty about origins or to the particular cues provided by the social setting at different times [51]. Generational differences in knowledge about ethnic origins were confirmed by research conducted by the U.S. Bureau of the Census on the 1980 decennial census [52]. Based on her research, Waters found that ethnic identity was constructed by using knowledge about ancestries in a person’s background. “This information is selectively used in the social construction of ethnic identification within the prevailing historical, structural, and personal constraints.” In other words, she wrote, “We pick and choose our ancestry(ies), and ‘lose’ some of our background; we exercise our ethnic options” [53].

## **“OTHER” RACE CATEGORY**

Racial and ethnic identification are also highly sensitive to the method of data collection (self-identification or observation by a third person), including terminology, question order in the instrument, and whether identification is made by the respondent or the observer. Conceptual difficulties are further exacerbated by ambiguity in question wording or by the requirement that respondents select only one race category [54].

One of the most serious and long-standing problems in deriving the estimates for racial and ethnic group statistics has been the residual “Other Race” category by respondents who either choose not to self-identify or find other categories inappropriate. Research has consistently shown that Hispanics constitute the largest number of respondents in this category [55]. People who self-identify as “multiracial” may also be relegated to this category. Experiments in questionnaire design with a multiracial population found that including a multiracial category in the list of races significantly improved the accuracy of racial classification by multiracial subjects [56]. Support for the “social construction of identity” that was asserted by the American Anthropological Association, was also found in comments made by these multiracial subjects, some of whom said that they felt “marginalized” by the residual “Other” category [57].

The 1980 decennial census (long form) introduced a new question on “ancestry” to obtain information about ethnicity. Levin and colleagues compared ancestry and race questions in the 1980 enumeration and sample [58]. They found a significant discrepancy in the number of people who self-identified as American Indian (race item) and also reported American Indian ancestry: 1.5 million and more than 6.7 million, respectively. Research conducted on the 1980 census by Snip also reinforced the discrepancy in self-identification and reported American Indian ancestry [59]. Snip noted that the “or” in the instructions for the question about personal or ancestral heritage created ambiguity for American Indians, Eskimos, and Aleuts. The result was that “ethnicity could have significantly different meanings, depending on whether respondents report their own ethnic identity or the ethnicity of their ancestors, among whom there may be many different ethnic groups represented” [60]. Because American Indians could check off only one race but were permitted multiple entries for ancestry, the result was dramatically different estimates of the American Indian population. The magnitude of the differences was so great that it was “difficult to reconcile the differences” in the estimates. “Very clearly,” he went on to say, “the race and ethnic ancestry questions reveal different populations” [61].

The 1996 Race and Ethnic Targeted Test conducted in 1996 by the U.S. Bureau of the Census as part of the review of Statistical Policy Directive 15 and preparations for the Year 2000 census, reinforced the finding that permitting multiple responses for racial identification would alter significantly and differentially the counts for different populations. American Indians, Alaska Natives, and Asian and Pacific Islanders, who experience higher rates of intermarriage compared to whites, were much more affected by the “Select more than one” option than the white and African American populations [62].

## **QUESTIONNAIRE “FRAME OF REFERENCE”**

Question sequencing order also creates a “frame of reference” for the respondent. Martin and colleagues conducted a split-ballot experiment in 1987, as part of their research for the 1990 decennial census long form survey, which was designed to reduce

the number of nonresponses and “Other race” reporting by Hispanics [63]. They found that the “race item set a frame of reference for the Hispanic origin item that accentuated the perceived redundancy of the [Hispanic origin] item” [64]. When the Hispanic origin item was placed immediately after race, the Hispanic origin item nonresponse rates ranged from 20 to 30 percent, but when placed three items after race, the nonresponse rates fell to between 13 and 17 percent. Although the goal of reducing the reporting of “Other race” by Hispanics was met for native-born, it was not for the immigrant sample population. Once again, there appeared to be empirical evidence that ethnicity was culturally constructed. The researchers found observable differences between populations depending on their status as immigrant or native-born, the former not yet socialized to American conceptions of ethnicity or race. Similar question order effects were also in the large scale 1995 Race and Ethnic Targeted Test (RAETT) mail-back survey conducted by the U.S. Bureau of the Census.

The cognitive research conducted by McKay and de la Puente also found that question order reinforced whatever conceptual confusion the respondent may have experienced. There was, for example, no confusion if the multiracial question (adding a category to the list of races) preceded the question on ancestry. This was not the situation, however, when the multiracial question category followed the question on the respondent’s ancestry or ethnic origin. Recalling a remote ancestor of a different race on the ancestry question led respondents to “think aloud about whether they should say they were multiracial, in light of how they had answered the ancestry question” [65]. The laboratory study of multiracial women conducted by the University of Illinois research team also found that the respondents were highly sensitive to the question form. Self-identification as a “multiracial” person was dramatically different when this category was offered for the racial classification question [66].

### **SELF- AND OBSERVER IDENTIFICATION**

Identifying race and ethnic origin are also influenced by who provides the information. Self- and observer identification of a subject’s race and ethnic origin have been found to differ significantly. Consistency and reliability of racial or ethnic identity vary by racial and ethnic groups, as well. Evaluations of vital statistics records have revealed that racial and ethnic identity varies over the life course, as a function of whether individuals are classified in birth and death records by self or third person observers [67]. Hahn and colleagues found that ethnic identity could change as many as four times over the life course [68]. Hahn and colleagues also found considerable inconsistencies in racial and ethnic classification in vital statistics records for infants [69]. Comparing funeral directors’ and interviewers’ assignment of race with self-identification by ancestry (race or ethnicity) indicated high consistency for “whites” and “blacks,” but low consistency for other races or ethnic groups. The researchers concluded that this inconsistency had significant effects on the estimates of birth and mortality rates.

The analysis by Hahn and colleagues of the status of federal health statistics that depend on the reporting by third parties, was a broadside attack on the assumptions underlying the collection of racial and ethnic group information. Obtaining “compatible, consistent, and exchangeable” data across agencies, a principal objective of Statistical Policy Directive 15, could not be ensured. A considerable body of evidence produced by the statistical agencies themselves would also support their assessment. Studies indicated that the reported statistics had to be carefully assessed for their quality because agencies did not verify the accuracy of the racial and ethnic data or accurately designate

nationality groups in a correct category [70]. Totals across surveys varied considerably [71]. Even within the same agency, such as the U.S. Bureau of the Census, terminology of racial and ethnic group categories varied [72].

### CONCLUDING REMARKS

Race has structured the organization of the state and civil society, “shaping social identities and institutions in significant ways” [73]. The politics of race and ethnicity have been laden with the competing claims of social groups. The enumeration of race has always reflected, as sociologist Waters noted in her testimony before Congress in 1997, the “enormous importance of the black/white color line in our society and the distinctive legacy of slavery” [74]. Racial classification has historically reproduced the social structure and the state’s conception of majority-minority status. The official category domains and nouns applied to racial and ethnic characteristics have implicitly reinforced a contested terrain of opposing and contrary claims about the autonomous pursuit of one’s “plan of life,” membership in the polity and social group, and promotion of the general welfare, which have resonated throughout U.S. history [75].

Racial and ethnic statistics produced by federal agencies have far-reaching effects for governments, business, and scientific bodies. Race and ethnicity statistics are key to implementing a very large number of federal, state, and municipal laws and regulations, vital to making policy decisions, employed in numerous programs, and are a critical factor in the basic research behind numerous policies [76]. The numbers affect how programs are administered, decisions about political representation and intergovernmental transfers, how samples based on counts supplied by the decennial census are drawn, and the interpretations of trends over time. The theoretical, conceptual, definitional and measurement problems inherent in the classification of racial and ethnic groups are, however, significant. These problems derive in part from their interaction with public policy that represents the traces and artifacts of historical controversy and dispute. The statistics resonate with the large-scale social and economic structural transformations and the larger policy debates taking place in the political system. The statistics embody socially meaningful biases. What the empirical evidence collected over the past 20 years demonstrates most powerfully, however, is that the conceptual basis and the subjective meaning of race and ethnicity are essential for understanding why racial and ethnic statistics will always maintain their problematic status.

Perhaps the most important aspect of any understanding of the U.S. federal statistical system is its complexity. Government information providers should expect difficulty in evaluating the quality or accuracy of the numbers related to statistics on race and ethnicity. Statistical data cannot be expected to resolve the great problems that divide a nation; they can, however, illuminate the policy debates about what must be done. These statistics will always be, as Stratford notes, “an imperfect representation of reality,” and, as such, must be interpreted with great care by the government information specialist and researcher [77]. This understanding of the intellectual task that confronts the government documents specialist is the foundation for providing high quality services.

### NOTES

1. Juri Stratford, “Responding to Reference Queries for Numeric Data and the Problems Inherent in Interpreting Statistical Sources: A Note,” *Journal of Government Information* 25, no. 5 (1998):413–417. The same problem is identified in the analysis of ethnic achievement of Polish American women. See: Thomas Duszak, “Polish American Women: Tracing Ethnic Achievement in Published U.S. Government Sources,” *Journal of Government Information* 25, no. 1 (1998):47–71.

2. This article is based on the author's current research on the meaning of the political discourse about the racial and ethnic group statistics in the government review to modify the classification system for racial and ethnic data, "Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity." The original standard was issued as Revised Exhibit F to Office of Management and Budget (OMB) Circular No. A-46 in 1977, and became widely known as "Statistical Policy Directive 15." This name was acquired while the statistical policy function was at the Department of Commerce in the Office of Federal Statistical Policy and Standards, whose function was later transferred to the OMB's Office of Information and Regulatory Affairs. For the purpose of this article, the author continues to refer to its widely known name of "Statistical Policy Directive 15," although the name change took effect during 1998.
3. Current terminology for the labels applied to race and ethnic group is employed except where historically accurate. Please note that double quotes are placed around the nouns associated with racial and ethnic identity in order to emphasize their socially constructed status. The naming conventions applied by the Bureau of the Census reflected space and formatting considerations, but, more importantly, the nouns reflected common usage and referents by official record keepers of the social group's status in the society at the time the decennial census was conducted.
4. See, as one example, the relationship between the census and development of programs and policies: U.S. Congress, House of Representatives, Committee on Post Office and Civil Service, *Effect of Census Statistics on Federal Programs and Reporting Requirements: Hearings before the Subcommittee on Census and Statistics*, 93rd Cong., 2nd sess., 31 May, 1 June 1974.
5. American Anthropological Association, *Response to OMB Directive 15* [September 1997] available from <<http://www.ameranthassn.org/comsumm.htm>>.
6. The ordering of the race categories in every census schedule between 1790 and 1990 illustrates this. "White" is the first racial noun identified on the list of races.
7. Stephen H. Riggins, ed., *The Language and Politics of Exclusion: Others in Discourse* (Thousand Oaks, CA: Sage Publications, Inc., 1997); Clara Rodriguez and Hector Cordero-Guzman, "Placing Race in Context," *Ethnic and Racial Studies* 15, no. 4 (1992):523-542.
8. Race and ethnic group statistics have also been collected as part of other federal data collections, including administrative records and surveys, since the turn of the century. For a history of the federal statistical system, see Joseph W. Duncan and William C. Shelton, *Revolution in United States Government Statistics, 1926-1976* (Washington, DC: Office of Federal Statistical Policy and Standards, 1978).
9. Since 1900, the U.S. Bureau of the Census has employed 26 different nouns to identify the population. Only "White," "Chinese," and "Japanese" appear on all twentieth-century enumeration schedules. The categories have been based on geographical location, national origin, language, minority status, and physical characteristics. For more discussion about the variation, see: American Anthropological Association, *Press Release/OMB 15* (Washington, DC: American Anthropological Association, 1997); Reynolds Farley, "Questions about Race, Spanish Origin and Ancestry: Findings from the Census of 1990 and Proposals for the Census of 2000," in *1991 Proceedings of the Social Statistics Section* (Arlington, VA: American Statistical Association, 1991), 356-361. The preeminent social history of the census is Margo Anderson, *The American Census: A Social History* (New Haven, CT: Yale University Press, 1988).
10. David E. Hayes-Bautista and Jorge Chapa, "Latino Terminology: Conceptual Bases for Standardized Terminology," *American Journal of Public Health* 7, no. 1 (1987):61-68.
11. This is a point that was made over and over again in the testimony presented by Native Hawaiians and American Indians during congressional and OMB hearings and in the hundreds of letters received by OMB following the publication of three *Federal Register* notices and request for public comment. See U.S. Office of Management and Budget, *Public Hearings in the Matter of Standards for the Classification of Federal Data on Race and Ethnicity*, Boston, 7 July 1994; Denver, 11 July 1994; San Francisco, 14 July 1994; Honolulu, 18-19 July 1984 (mimeo).
12. U.S. Department of Commerce, Bureau of the Census, *200 Years of U.S. Census Taking: Population and Housing Questions, 1790-1990* (Washington, DC, 1989). Henceforth, referred to as Bureau of the Census, *200 Years*.
13. Noel Ignatiev, *How the Irish Became White* (New York: Routledge, 1995).
14. Bureau of the Census, *200 Years*, 83.
15. In 1930, separate figures were obtained for Mexicans, defined as "all persons born in Mexico, or having parents born in Mexico, who were not definitely White, Negro, Indian, Chinese, or Japanese, would be returned as Mexicans." In the 1960 and 1970 censuses, "White persons of Spanish surname" in the five states of the "Mexican" Southwest, were classified as "Mexicano." In the 1990 census, Item 7 ("Hispanic/Spanish Origin") established an "Other Spanish/Hispanic" category that named a long set of examples, concluding with the words "and so on."
16. Bureau of the Census, *200 Years*, 26.
17. Bureau of the Census, *200 Years*, 36.
18. Bureau of the Census, *200 Years*, 41.
19. Bureau of the Census, *200 Years*, 50.
20. Bureau of the Census, *200 Years*, 60.

21. Bureau of the Census, *200 Years*, 60.
22. Bureau of the Census, *200 Years*, 47.
23. Bureau of the Census, *200 Years*, 57.
24. Bureau of the Census, *200 Years*, 60.
25. Bureau of the Census, *200 Years*, 83.
26. Bureau of the Census, *200 Years*, 89.
27. Titles of Bureau of the Census and other agency publications reinforced the differentiation of the “other” as members of the non-white population. Until recent decades, publications on the racial composition of the nation were often entitled, “Whites and All Others.”
28. Bureau of the Census, *200 Years*, 78.
29. Bureau of the Census, *200 Years*, 89.
30. “Slave Descendant Fights Race Listing,” *New York Times*, 15 September 1982; F. F. Marcus, “Louisiana Repeals Black Blood Law,” *New York Times*, 6 July 1983.
31. See: *Doe v. State*, 479 So. 2d La. 369 (1985) and the subsequent appeal: *Doe v. State*, 485 So. 2d La. 60 (La. Ct. App. 1986).
32. U.S. Federal Intergency Committee on Education, *Report of the Ad Hoc Committee on Racial and Ethnic Definitions* (Washington, DC, 1974); U.S. Office of Management and Budget, “Standards for the Classification of Race and Ethnicity,” *Federal Register* 59 (1 June 1994): 29831-35; U.S. Office of Management and Budget, “Standards for the Classification of Federal Data on Race and Ethnicity,” *Federal Register* 60 (28 August 1995): 44673-93; U.S. Office of Management and Budget, “Recommendations from the Intergency Committee for the Review of the Racial and Ethnic Standards to the Office Management and Budget Concerning Changes to the Standards for the Classification of Federal Data on Race and Ethnicity,” *Federal Register* 62 (9 July 1997): 36874-946; U.S. Congress, House of Representatives, Committee on Post Office and Civil Service, *Review of Federal Measurement of Race and Ethnicity: Hearings before the Subcommittee on Census, Statistics and Postal Personnel*, 103rd Cong., 1st sess., 14 April, 30 June, 29 July, 3 November 1993; U.S. Congress, House of Representatives, Committee on Government Reform and Oversight, *Federal Measures of Race and Ethnicity and the Implications for the 2000 Census: Hearings before the Subcommittee on Government Management, Information, and Technology*, 105th Cong., 1st sess., 23 April, 22 May, 25 July 1997; U.S. Office of Management and Budget, *Public Hearings*.
33. Mid-decade amendments to the Immigration and Nationality Act of 1952 and termination of the Bracero Program during the 1960s altered the demography of the nation and “set the stage for two of the major shifts in the character of immigration that were to occur in subsequent decades” (Barry Edmonston and Jeffrey N. Passel, eds., *Immigration and Ethnicity: The Integration of America’s Newest Arrivals* (Washington, DC: Urban Institute Press, 1994), 6.
34. Yen L. Espiritu, *Asian American Panethnicity: Bridging Institutions and Identities* (Philadelphia, PA: Temple University Press, 1992); Gareth Davies, *From Opportunity to Entitlement: The Transformation and Decline of Great Society Liberalism* (Lawrence, KS: University Press of Kansas, 1996).
35. See, for example, U.S. Congress, House of Representatives, Committee on Post Office and Civil Service, *Accuracy of 1970 Census Enumeration and Related Matters: Hearings before the Subcommittee on Census and Population*, 91st Cong., 2nd sess., 15, 16, 22–24, 29, 30 September 1970; U.S. Congress, House of Representatives, Committee on Post Office and Civil Service, *Economic and Social Statistics for Americans of Spanish Origin: Hearing before the Subcommittee on Census and Population*, 94th Cong., 1st sess., 21 March 1975; U.S. Congress, House of Representatives, Committee on Post Office and Civil Service, *Equal Opportunity Hearing in Denver, Colorado: Hearing before the Subcommittee on Postal Facilities, Mail, and Labor Management*, 94th Cong., 1st sess., 21 May 1969; U.S. Congress, Senate, Committee on Labor and Human Resources, *Migrant and Seasonal Farmworker Powerlessness, Who is Responsible: Hearings before the Subcommittee on Migratory Labor*, 91st Cong., 2nd sess., 1, 20, 21, 24 July 1970; U.S. Congress, House of Representatives, Committee on the Judiciary, *Federal Government’s Role in the Achievement of Equal Opportunity in Housing: Hearings before Subcommittee No. 4*, 92nd Cong., 1st sess., 27 October, 2, 3, 10, 11 November, 2, 8, 9 December 1971; and 2nd sess., 15 June 1972; U.S. Congress, House of Representatives, Committee on Post Office and Civil Service, *1980 Census: Hearings before the Subcommittee on Census and Population*, 94th Cong., 2nd sess., 1-2 June 1980; U.S. Congress, House of Representatives, Select Committee on Immigration, *Immigration in the U.S.*, 95th Cong., 2nd sess., 4–7 April 1978.
36. “Voice” should be understood as (1) the appearance of entries on administrative forms and other recording instruments, which by their very appearance as an item on a form acknowledges a particular identity or social status; and (2) participation in an administrative process by which these entries are developed.
37. Ruth B. McKay and Manuel de la Puente, “Cognitive Research in Designing the CPS Supplement on Race and Ethnicity” (paper presented at the U.S. Bureau of the Census 1995 Annual Research Conference, March 19–22, 1995, Rosslyn, VA). Henceforth, referred to as “McKay and de la Puente, Cognitive Research.” Although the McKay and de la Puente study was among the first extensive governmental research to probe explicitly the meaning of the concepts, their findings only reinforced other research conducted by members of the U.S. Bureau of the Census that had investigated the reporting of the “race”

- item in the decennial censuses. For example, the "Content Reinterview" surveys that follow the fielding of a decennial census have always found a misreporting of race, with people self-identifying as "Irish," "Polish," "German," or "Italian." See, for example, Kathryn F. Thomas and Tamara L. Dingbaum, "How Good are These Data Anyway? The 1990 Content Reinterview Survey," in *1992 Proceedings of the Section on Survey Research Methods* (Alexandria, VA: American Statistical Association, 1992), 369–374.
38. U.S. Department of Labor, Bureau of Labor Statistics, *Testing Methods of Collecting Racial and Ethnic Information: Results of the Current Population Survey Supplement on Race and Ethnicity* (BLS Statistical Notes 40) (Washington, DC, 1996).
  39. Respondents were asked, "Please tell me what you think is the most important characteristic that defines race?"; and "Do you think there is any difference between race, ethnicity, and ancestry?" (McKay and de la Puente, "Cognitive Research," 5).
  40. McKay and de la Puente, 4, 7.
  41. Q4a asked, "You selected [Fill: race] from the list I read to you. Do you also have a more specific group within [Fill: race] that you belong to?" and "If yes," Q4b asked, "What is the name of the specific group?" McKay and de la Puente, 4, 7.
  42. First, the researchers asked whether the person was of Hispanic, Latino, or Spanish origin. Question two followed, "People sometimes think of customs, or language, or physical appearance, or country of birth when they think of race. What comes to your mind as most important when you think of a person's race?" Question three asked, "People sometimes think of customs, or language, or physical appearance, or country of birth when they think of race. What comes to your mind as most important when you think of race?" The next question asked, "And, finally, what comes to your mind as most important when you think of a person's ancestry?" McKay and de la Puente, 4, 7.
  43. See note 38.
  44. Timothy P. Johnson, Jared B. Jobe, Diane O'Rourke, Seymour Sudman, Richard Warnecke, Noel Chavez, Gloria Chapa-Resendez, and Patricia Golden, "Dimensions of Self Identification Among Multiracial and Multiethnic Respondents in Survey Interviews" (paper presented at the 1996 Annual Research Conference and Technology Interchange meeting, Arlington, VA, March 1996).
  45. McKay and de la Puente, "Cognitive Research," 6.
  46. Clara E. Rodriguez, *Puerto Ricans: Born in the U.S.A.* (Boulder, CO: Westview, 1991).
  47. Hilary L. Surratt and James A. Inciardi, "Unraveling the Concept of Race in Brazil: Issues for the Rio de Janeiro Cooperative Agreement Site," *Journal of Psychoactive Drugs* 30, no. 3 (1998):255–260; Clara E. Rodriguez, "Race, Culture, and Latino 'Otherness' in the 1980 Census," *Social Science Quarterly* 71, no. 4 (1991):930–937.
  48. Jorge Calerón, "'Hispanic' and 'Latino': The Viability of Categories for Panethnic Unity," *Latin American Perspectives* 19, no. 4 (1992):37–44; David Hayes-Bautista and Jorge Chapa, "Latino Terminology: Conceptual Bases for Standardized Terminology," *American Journal of Public Health* 77, no. 7 (1987):61–68; Johnson et al., *Dimensions of Self-Identification*, 1996; Suzanne Oboler, "The Politics of Labeling: Latino/a Cultural Identities of Self and Others," *Latin American Perspectives* 19, no. 4 (1992):18–36; Clara E. Rodriguez, "Racial Classification Among Puerto Rican Men and Women in New York," *Hispanic Journal of Behavioral Sciences* 12, no. 4 (1990):366–379; Rodriguez, 1991; Clara Rodriguez and Hector Cordero-Guzman, "Placing Race in Context," *Ethnic and Racial Studies*, 15, no. 4 (1990):523–542; Clara E. Rodriguez, Aida Castro, Oscar Garcia, and Analisa Torres, "Latino Racial Identity: In the Eye of the Beholder?," *Latino Studies Journal* (1991):33–48; Fernando M. Trevio, "Standardized Terminology for Hispanic Populations," *American Journal of Public Health* 77, no. 1 (1987):69–72.
  49. Joan Nagel, "Constructing Ethnicity: Creating and Recreating Ethnic Identity and Culture," *Social Problems* 41 (1994):152–176.
  50. Charles E. Johnson, *Consistency of Reporting Ethnic Origin in the Current Population Survey* (Bureau of the Census Technical Paper 31) (Washington, DC, 1974), 6–7. Low consistency groups were those who reported "Other Spanish" origin, Central or South American origin, and those who reported that they did not know their ethnic origin. Shifts and inconsistencies between the two surveys also occurred for Welsh, Scottish, and English ethnic origins, which comprise the oldest ethnic groups in the United States and who have had high rates of intermarriage. Henceforth, referred to as *Consistency of Reporting*.
  51. Johnson, *Consistency of Reporting*, 1974; Richard D. Alba and Mark B. Chamlin, "A Preliminary Examination of Ethnic Identification among Whites," *American Sociological Review* 48, no. 2 (1983):240–247; Mary Waters, *Ethnic Options: Choosing Identities in America* (Berkeley, CA: University of California Press, 1990).
  52. Nampeo R. McKenney, Reynolds Farley, and Michael J. Levin, "Direct and Indirect Measures of Ethnicity: How Different Definitions Affect the Size and Characteristics of Various Ethnic Groups," in *1983 Proceedings of the Social Statistics Section* (Washington, DC: American Statistical Association, 1983), 123–130.
  53. Waters, *Ethnic Options*, 20.
  54. This rule was altered by the revision of the standards for classifying racial and ethnic data. There is not, as yet, a sufficiently large body of empirical evidence to indicate the magnitude of the effects of the rule change. The small amount of evidence does, however, reveal significant effects for selected population groups, as discussed below.
  55. Manuel de la Puente, "A Multivariate Analysis of the Census Omission of Hispanics and Non-Hispanic Whites, Blacks, Asians, and American Indians: Evidence from Small Area Ethnographic Studies," in

- 1993 *Proceedings of the Section on Survey Research Methods* (Alexandria, VA: American Statistical Association, 1993), 641–646; Johnson, *Consistency of Reporting Ethnic Origin*, 1974; Eduardo W. Fernandez and Nampeo R. McKenney, “Identification of the Hispanic Population: A Review of Census Bureau Experiences,” in *1980 Proceedings of the Section on Survey Research Methods* (Washington, DC: American Statistical Association, 1980), 358–363.
56. Johnson et al., “Dimensions of Self Identification,” 1996.
  57. See, for example, F. James Davis, ed., *Who is Black?: One Nation’s Definition* (University Park, PA: Pennsylvania State University Press, 1991); Naomi Zack, ed., *American Mixed Race: The Culture of Microdiversity* (Lanham, MD: Rowman & Littlefield Publishers, 1995).
  58. Michael J. Levin, Nampeo R. McKenney, and Paul A. Berman, “Uses and Interpretations of Racial and Ethnic Data from the U.S. Census,” in *1984 Proceedings of the Social Statistics Section* (Washington, DC: American Statistical Association, 1984), 88–95.
  59. C. Mathew Snip, “Who are American Indians? Some Observations about the Perils and Pitfalls of Data for Race and Ethnicity,” *Population Research and Policy Review* 5 (1986):237–252.
  60. Snip, 240.
  61. Snip, 241.
  62. U.S. Department of Commerce, Bureau of the Census, *Results of the 1996 Race and Ethnic Targeted Test* (Population Division Working Paper Series) (Washington, DC, 1997).
  63. Elizabeth Martin, Teresa J. DeMaio, and Pamela C. Campanelli, “Context Effects for Census Measures of Race and Hispanic Origin,” *Public Opinion Quarterly* 54, no. 4 (1990):551–566.
  64. Martin, Demaio, and Campanelli, 551.
  65. McKay and de la Puente, “Cognitive Research,” 5.
  66. Johnson et al., “Dimensions of Self Identification,” 4–5. The researchers found that the introduction of a “multiracial” option dramatically increased the response rates, when compared to the options of “Other-specify” and “Other.”
  67. Robert A. Hahn, “The State of Federal Health Statistics on Racial and Ethnic Groups,” *Journal of the American Medical Association* 267, no. 2 (1992):268–271; Robert A. Hahn and Steven Eberhart, “Life Expectancy in Four U.S. Racial/Ethnic Populations: 1990,” *Epidemiology* 6, no. 4 (1995):350–355.
  68. Robert A. Hahn, Benedict I. Truman, and Nancy D. Barker, “Identifying Ancestry: The Reliability of Ancestral Identification in the United States by Self, Proxy, Interviewer, and Funeral Director,” *Epidemiology* 7, no. 1 (1995):75–80.
  69. Robert A. Hahn, Joseph Muolinare, and Steven M. Teutsch, “Inconsistencies in Coding of Race and Ethnicity Between Birth and Death in US Infants,” *Journal of the American Medical Association* 267, no. 2 (1992):259–263.
  70. Summaries of findings throughout the statistical agencies can be found in several U.S. General Accounting Office reports: *Federal Data Collection: Agencies’ Use of Consistent Race and Ethnic Definitions, Report to the Chairman, Committee on Government Operations, House of Representatives* (GAO/GDD-93-25) (Washington, DC, 1992); *Federal Data Collection: Measuring Race and Ethnicity is Complex and Controversial, Testimony before the Subcommittee on Census, Statistics, and Postal Personnel, Committee on Post Office and Civil Service, House of Representatives* (GAO/T-GGD-93-21) (Washington, DC, 1993); *Statistical Agencies: Collection and Reporting of Race and Ethnicity Data, Testimony before the Subcommittee on Governmental Reform and Oversight, House of Representatives* (GAO/T-GGD-97-92) (Washington, DC, 1997).
  71. A recent comparison of 1990 decennial census data, which rely on self-identification, and National Center for Health Statistics’ birth and death records, which rely on observer identification, supports earlier empirical evidence about consistency across agencies. Estimates differ in the way that race and Hispanic origin are assigned. See Larry Sink, *Race and Ethnicity Classification Consistency between the Census Bureau and the National Center for Health Statistics* (Technical Working Paper No. 17) (Washington, DC, February 1997) [4 November 1998]. Available from: <<http://www.census.gov/population/www/documentation/twps0017.html#abstract>>
  72. For example, the 1990 decennial census listed a category “American Indian or Alaskan Native,” but the National Crime Victimization Survey has coded Alaskan natives as “Aleut” or “Eskimo” on “Control Card Item 23.”
  73. Michael Omi and Howard Winant, *Racial Formation in the United States From the 1960s to the 1990s*, 2nd ed. (New York: Routledge, 1994), vii.
  74. Committee on Government Reform and Oversight, *Federal Measures*, 440.
  75. There is abundant literature on this subject. See, for example, Leobardo F. Estrada, “The Politics of the Census: A Reflection of the Dilemmas in U.S. Society,” in *Challenges of Measuring an Ethnic World: Science, Politics and Reality* (proceedings of the Joint Canada-United States Conference on the Measurement of Ethnicity, April 1–3, 1992) (Washington, DC, Commerce Department and Ottawa: Statistics Canada, 1993), 497–512; Anderson, *The American Census*; Barry Edmonston and Charles Schultze, *Modernizing the U.S. Census* (Washington, DC: National Academy Press, 1994).
  76. For an enumeration of the laws, see, for example: U.S. Department of Commerce, Bureau of the Census, *Preparing for the Census 2000, Subjects Planned for Census 2000: Federal Legislative and Program Uses* (Washington, DC: Bureau of the Census, May 1997).
  77. Stratford, “Responding to Reference Queries,” 416.