Race/Ethnicity Code Changes: Mapping/Bridging and Reporting Data

Texas Association for Institutional Research

Data Advisory Committee

Maryann Ruddock, chair   University of Texas at Austin
Kate Amorella   Texas Women's University
Cindy Dutschke   Texas A&M University
Kathy Friedrich   College of the Mainland
Sue Herring   Baylor University
Tom Martin   Collin County Community College District
Darline Morris   Texas State Tech College
Martha Oburn   Houston Community College System
Larry Redlinger   University of Texas at Dallas
Leona Urbish   Rice University

October 2009

Race/Ethnicity Code Changes:Mapping/Bridging and Reporting Data

All educational institutions (elementary and secondary schools, two- and four-year colleges and universities) that report data to the federal government are in the process of changing how they report race/ethnicity data for students and employees. The federal government first defined race/ethnicity categories in 1977. They have remained unchanged until changes to race/ethnicity code categories were mandated by the U.S. Office of Management and Budget in 1997, and first appeared in the 2000 Census collection. A transition period was given in order for agencies to review the results of the census collection. These reporting standards (categories) were finalized by the Department of Education (for students) and the Department of Labor (for employees). The October 19, 2007 Federal Register published "Final Guidance on Maintaining, Collecting, and Reporting Racial and Ethnic Data to the U.S. Department of Education". The new categories are required for reporting fall 2010 data.

Race/Ethnic Code Categories - Reporting of Data Internally
While the reporting of race/ethnicity data to the federal government is mandated by them, there is no mandate for how institutions use the data internally and how they choose to portray their ethnic composition. Most objections to the new race/ethnicity categories (after questioning why Hispanics were singled out) involved the reporting of any combination of the five race codes as "two or more". The Department of Education received many comments about the "loss" of information that occurs when respondents are able to identify combinations of the five race categories but the data are aggregated into a single "two or more" category for reporting to the federal government. The Department stands behind its guidelines.

It is up to individual institutions to determine how they wish to report data not going to the federal or state governments. Since data are collected and need to be stored by individual response (unit record level), the institution may wish to report race/ethnicity data by a combination of categories. The federal government considered, then rejected, a min-max reporting (e.g., black only versus black and any other categories), but institutions may want to consider this for internal reporting. While the federal government is not asking for race data in combination with Hispanic ethnicity, institutions have the capability to report Hispanic plus another race or races.

The State of Texas has also mandated reporting of the new race/ethnicity codes. State reporting follows federal reporting, except for the "two or more" category. For state reporting, the first "two or more races" category consists of "two or more races with at least one being Black" and the second "two or more races" category is labeled "two or more races" and includes all other "two or more races".

The issue of handling data reporting outside of the federal and state definitions also extends to looking at trend data. The federal government (http://www.whitehouse.gov/omb/inforeg/re_app-ctables.pdf) states that the most important criterion for looking forward (mapping the "current" categories to the "new") is "congruence with respect to respondent's choice", while the most important criterion for bridging to the past is "measuring change over time". Other considerations include range of applicability, meet confidentiality and reliability standards, statistically defensible, ease of use, skill required, understandability and communicability, and congruence with respondent's choice. For congruence with respondent's choice, the underlying logic of the tabulation procedures must reflect to the greatest extent possible the full detail of racial reporting (given that there were changes in the instructions and the categories).

Mapping

Mapping of data generally refers to comparing old/current data to new data. The five current race/ethnicity categories are mapped to the seven "new" categories (international and unknown remain the same).

Old/Current (1979)	New (1997)
Hispanic	Hispanic (regardless of race)
American Indian or Alaska Native	American Indian or Alaska Native only
Black, non-Hispanic	Black only
White, non-Hispanic	White only
Asian or Pacific Islander	Asian only
	Native Hawaiian/Other Pacific Islander only
Nonresident Alien	Nonresident Alien
Unknown	Unknown
	Two or More

Mapping is generally used for placing current students and employees into new categories to show where they would be categorized if they do not re-identify. As seen in the table, there is no category that maps to Native Hawaiian/Other Pacific Islander only, nor to "two or more races".

Bridging

The federal government has conducted research on taking the "new" categories (referred to by the federal government as the 1997 codes) and bridging them back to the "old/current" categories (referred to as the 1979 categories) (http://www.whitehouse.gov/omb/inforeg/re_app-ctables.pdf).

New (1997)	Old/Current (1979)
Hispanic (regardless of race)	Hispanic
American Indian or Alaska Native only	American Indian or Alaska Native
Black only	Black, non-Hispanic
White only	White, non-Hispanic
Asian only	Asian or Pacific Islander
Native Hawaiian/Other Pacific Islander only	Asian or Pacific Islander
Nonresident Alien	Nonresident Alien
Unknown	Unknown
Two or More	['unknown' or use bridging methodology]

Bridging is generally used for reporting trend data, to measure change over time. The ideal bridging method, under this criterion, would be one that matches how the respondent would have responded under the 1979 standards had that been possible. In this ideal situation, differences between the new distribution and the old distribution would reflect true change in the distribution itself. (http://www.whitehouse.gov/omb/inforeg/re_app-ctables.pdf)

When bridging back from the "new" categories to the "old/current" categories, Hispanic, American Indian, White, and Black correspond fairly well, given that for the new categories Hispanic trumps all races. Asian only and Native Hawaiian/Other Pacific Islander would correspond to Asian or Pacific Islander. Since only 0.1% (2,203) of the 18 – 24 year-old Texan respondents to the 2000 Census responded that they were Native Hawaiian/Other Pacific Islander, there is not much data lost by combining these categories. The area where most data is being lost is in the "two or more races" category. All of the "two or more races" responses could be assigned to "unknown", but that would result in the loss of information.

According to the federal government (Federal Register: October 19, 2007, Volume 72, Number 202, Page 59266-59279):

> States, educational institutions, and other recipients also may propose to "bridge'' the "two or more races'' category into single race categories or the new single race categories into the previous single race categories. Bridging involves adopting a method for being able to link the new data collected using the two-part question with data collected before the publication of this guidance by the Department. If States, educational institutions, and other recipients do bridge data, the bridging method should be documented and available for the Department to review, if necessary.

One method is to redistribute the new data collected under this guidance using the new racial and ethnic categories and relate them back to the racial and ethnic categories used before the publication of this guidance. For example, if a State's new data collection results in 200 students falling in the "two or more races'' category at the same time that there is a combined drop in the number in the two single race categories of Black or African American students and White students, the State can adopt a method to link the 200 students in the "two or more races'' category to the previously used Black and White categories.

Another method is assigning a proportion of the "two or more races'' respondents into the new five single-race categories. If educational institutions or other recipients choose to bridge, they may use one of several bridging techniques. For example, they may select one of the bridging techniques in OMB's Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity (http://www.whitehouse.gov/omb/inforeg/statpolicy.html#dr).

Educational institutions and other recipients also may choose to use the allocation rules developed by OMB in its Guidance on Aggregation and Allocation of Data on Race for Use in Civil Rights Monitoring and Enforcement (http://www.whitehouse.gov/omb/bulletins/b00-02.html). If a bridging technique is adopted, the same bridging technique must be used when reporting data throughout the educational institution or other recipient. For example, the same bridging technique should be used by the entire State for the purposes of NCLB.

The different bridging methodologies allocate multiple race/ethnicity responses to single race categories in different ways. Since single race responses match exactly and more than two responses are a small percentage of multiple responses, the two-choices responses have been chosen to show the difference in how the different bridging methodologies allocate. **Table A** presents how data would be reported, that is, what category the multiple response would be allocated to, for a number of methodologies that have shown up in the literature. Each bridging methodology is compared to the mandated federal and state reporting guidelines.

Race/Ethnicity Bridging Methodologies - How Data Would Be Reported

Collection Response	Federal Reporting	State Reporting	Office of Management and Budget						Texas State Demographer Alternative IV	All Inclusive		Consortium on Financing Higher Education	Federal Civil Rights Reporting (not for bridging)
			Deterministic Whole Assignment			Deterministic Fractional Assignment							
			Smallest Group	Largest Group other than White	Largest Group	Equal Fractions		Weighted Fractions					
White & Am Indian*	Two or More	Two or More	American Indian	Am Indian	White	.5 White	.5 American Indian	NA	White	White	American Indian	White	American Indian
White & Asian*	Two or More	Two or More	Asian	Asian	White	.5 White	.5 Asian	NA	White	White	Asian	Asian	Asian
Black & White*	Two or More	Two or More with one being Black	Black	Black	White	.5 Black	.5 White	NA	Black	Black	White	Black	Black
Am Indian & Black*	Two or More	Two or More with one being Black	American Indian	Black	Black	.5 American Indian	.5 Black	NA	Black	American Indian	Black	Black	allocate to the race the complaint was based on
Black & Hispanic	Hispanic	Hispanic	Black	Hispanic	Hispanic	.5 Black	.5 Hispanic	NA	Black	Black	Hispanic	Black	allocate to the race the complaint was based on
Black & Asian	Two or More	Two or More with one being Black	Black	Asian	Asian	.5 Black	.5 Asian	NA	Black	Black	Asian	Black	allocate to the race the complaint was based on
Am Indian & Hispanic	Hispanic	Hispanic	American Indian	Hispanic	Hispanic	.5 American Indian	.5 Hispanic	NA	Other	American Indian	Hispanic	Hispanic	allocate to the race the complaint was based on
Native Hawaiian & Asian	Two or More	Two or More	Native Hawaiian	Asian	Asian	.5 Native Hawaiian	.5 Asian	NA	Other	Native Hawaiian	Asian	Asian	allocate to the race the complaint was based on
Native Hawaiian & White	Two or More	Two or More	Native Hawaiian	Native Hawaiian	White	.5 Native Hawaiian	.5 White	NA	White	Native Hawaiian	White	Asian	Native Hawaiian

* identified by the federal government as most common double race combinations, used for Civil Rights monitoring and enforcement

Bridging Methodologies- Federal Office of Management and Budget
There are two basic ways of assigning multiple categories to a single category: whole assignment, where an individual's responses are assigned to a single racial category, and fractional assignment, where the individual's response is assigned to multiple categories.   Whole and fractional assignments can be made based on a set of deterministic rules or based on some probabilistic distribution.

On December 15, 2000, the U.S. Office of Management and Budget issued "Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity". The 1997 standards are the ones that are just now being implemented for all educational institutions. The discussion of bridging methodologies (Appendix A) is taken from that report. In addition, a mock data set (Table B) was devised for 1,000 hypothetical students, assigned to the 64 possible categories. These data were then used to show how the different bridging methodologies produce different results.

OMB Bridging Methodologies		
Assigned by a fixed rule or by a probability method?	Assigned to one or more than one category?	Methodology
Deterministic *A person is assigned to a category following a set of predetermined rules*	Whole *Person is assigned completely to one category*	Smallest Group
		Largest Group Other than White
		Largest Group
		Plurality
	Fraction *Person assigned partially to each selected group*	Deterministic Equal Fractions
		Deterministic Weighted Fractions
Probabilistic *A person is assigned to a category based on a probability distribution*	Whole *Person is assigned completely to one category*	Probabilistic Equal Fractions
		Probabilistic Weighted Fractions
		Hot Deck Imputation
	Fraction *Person assigned partially to each selected group*	*Not applicable*

Bridging Methodologies - State of Texas
The Office of the State Demographer for Texas has prepared a paper on "Comparing Race/Ethnicity Between the 2000 Census and Earlier Censuses" (http://txsdc.utsa.edu/txdata/redistrict/re-report.php). This paper presents assumptions used to bridge data from the 2000 Census (the "new" categories) back to the 1990 census categories (the "old/current" categories). However, concerning this methodology, it must be noted that the 2000 Census allowed respondents to fill in "some other race", an option not allowed for the IPEDS or Department of Labor reporting.

Previous work on race and ethnicity conducted by the Texas State Data Center and the Office of the State Demographer have concentrated on four large mutually exclusive race/ethnicity groups – Anglos (non-Hispanic whites), Blacks (non-Hispanic Blacks), Hispanics (of all races), and an Other category which consists of non-Hispanic persons from all other racial categories, with a majority of this group consisting of Asians. For the 2000 Census data, they continue to look at these categories.

The next step is to make specific allocation assumptions about certain race/ethnicity groups. One principle followed by the State Data Center is that "the allocation of parts of multiple race groups to different single race groups leads to allocation difficulties that are not easily resolved as one attempts to examine data on demographic or socioeconomic characteristics for the groups selected (e.g., there are extensive difficulties in using such allocation schemes with age-sex groupings and other demographic and socioeconomic data)." Therefore, there will be no partial allocations in the State Data Center methodologies.
1.   The Hispanic population will be considered a single group, regardless of the race of the respondent. For the 2000 Census, the three categories of White and Some Other Race, alone and in combination account for 97.7 percent of all Hispanics. Since Hispanics rarely identified a race other than White or Some Other Race, they are considered as a single group.
2.   The population of non-Hispanic persons indicating three or more racial groups represents only 0.093 percent of the non-Hispanic population (and only 0.09 percent of the total population). Because of this small number and the difficulty with partial allocations used for other purposes, the State Data Center allocates persons in all these three or more racial group categories of non-Hispanics to the Other category.
3.   Persons responding to a single race group are allocates as: non-Hispanic Whites are allocated to Anglo; non-Hispanic Blacks to the Black category; and non-Hispanic persons of American Indian and Alaska Native, Asian, Native Hawaiian and Other Pacific Islander, or Some Other Race to the Other category.
4.   The last group of respondents indicated two racial categories. They make up only 2.4 percent of Texas residents. There are four alternatives presented here for the allocation of two-race combinations on non-Hispanics to one of the four groups (Anglo, Black, Hispanic or Other). All

four alternatives have the basic assumptions that Hispanics of all races can be used as a single group; only non-Hispanic persons indicating a race of White alone are categorized as Anglo; only non-Hispanic persons indicating a race of Black alone are categorized as Black; and non-Hispanics of all other single and multiple racial groups alone or in combination can be categorized as Other.

Alternative I requires no assumptions about multi-racial groups except the assumption that Hispanics of all races can be included as a group.

Alternative II allocates according to the following: persons in the category of non-Hispanic White and Some Other Race were considered to be Anglo; persons who are non-Hispanic Black and Some Other Race are categorized as Black; and non-Hispanic persons in the American Indian or Alaska Native, Asian, Native Hawaiian and Pacific Islander, and Some Other Race were placed in the Other racial/ethnic group category. All other two and three race combinations of non-Hispanics are also placed in the Other non-Hispanic category.

Alternative III uses identical assumptions to those in Alternative II with one exception. The two-race combination of White/Black is allocated from the Other category to the Black category. The assumption here is that Black is a dominant identification relative to White in American Society.

Alternative IV assumes that within two-race categories that include Whites or Blacks, White and Black are dominant in terms of identification and that Black identification is dominant in the White/Black combination.

The State Data Center research finds Alternative IV the most reasonable for all groups. They have "also examined the four alternatives relative to their variation from expected historical patterns and Alternative IV again appears to be the most reasonable."

An additional way to present data that reflects the complexity of reporting on race is to use categories that may not be mutually exclusive. One example of this is categories that are "**all inclusive**". In this case distributions would be formed for each individual racial or ethnic group. Appendix F presents data for race/ethnicity and any combination, plus race/ethnicity only and any combination. The difference between these two groups is the race/ethnicity only count. This data display would answer the question "How many people on campus consider themselves to be Black/White/Asian/etc.?". The OMB report separates the racial categories from the ethnicity (Hispanic) category, just reporting all inclusive data for races.

Bridging Methodologies- Consortium on Financing Higher Education
Work has been done in the Institutional Research profession on bridging

methodologies and presented at professional conferences. One example of this is "Racial Identity and Government Classification: A Better Solution" presented by Anthony Broh and Stephen Minicucci at the 2008 forum of the Association of Institutional Research. This paper outlines current efforts from the College Board, Common Application, and the Consortium on Financing Higher Education (COFHE) to provide data collection instruments that comply with federal standards and meet the evolving needs of college admissions officers and administrators. In addition, the Consortium offers a bridging strategy for standardized reporting among schools that wish to produce trend lines.

According to Broh and Minicucci, the desirable characteristics for a bridging strategy would be:

Succinctness – The list of summary categories should be short, which may require folding very small groups into larger ones.

Comparability – The new summary categories should be comparable to data collected in a single-response format.

Verisimilitude – the summary should not marginalize the "multiples" into tabular footnotes, nor overstate minority populations by assuming all "multiples" are "minorities", nor overstate the majority population by assuming that "multiples" are not minorities.

Their classification rules are:

1. The classification applies only to US Citizens and Permanent Residents.
2. Respondents who answer the race/ethnicity question, but not the citizenship question are assumed to be US citizens or Permanent Residents; foreign includes those who responded to the citizenship question as "Foreign" (this applies to the Common Application for private institutions and reporting to the College Board)
3. Those not responding to either the first or second question are summarized as "Unknown"
4. Hawaiian and other Pacific Islander are classified as "Asian"
5. Those selecting more than two categories are classified as "more than two"
6. Black supersedes other identities; that is, a respondent choosing African American and *any other combination* of identities is summarized as Black
7. Hispanic supersedes identities other than Black
8. Asian supersedes White
9. American Indian identities do not supersede any other category

Which Bridging Methodology to Use?

One conclusion reached by the OMB study was that:

...the difficulty of forming a bridge to the past will differ depending on the particular racial group as reported under the 1977standards. Other racial groups also may be more or less likely to report multiple races in certain cases. For instance, the size of the population

reporting more than one race no doubt will differ by state, size of place, and also by some individual demographic characteristics such as the levels of income, education, and, especially, age. The various methods for creating the bridge could have different effects on the statistics for groups defined by these and other variables.

According to the State Demographer's Office:
> The four alternatives have different advantages and disadvantages. Alternative I requires no assumptions about multi-racial groups except the assumption that Hispanics of all races can be included as a group. Alternatives II-IV involve increasing amounts of allocation that allow the user to discern the level of allocation they believe feasible. Strictly for purposes of making comparisons between the four categories of Anglo, Black, Hispanic, and Other from 1990 to 2000, we believe that Alternative IV is to be recommended for most uses. This recommendation is based on the fact that it appears to provide values that are most reasonable relative to population change in the Anglo, Black, and Hispanic groups and within the Other group.

Reporting

The function of the presentation of race/ethnicity data is to answer a specific question. When respondents were only permitted to choose one race/ethnicity category out of a list of five, there was only one answer, regardless of the wording of the question. *How many Asian students are enrolled on your campus?* and *How many of you students consider themselves to be Asian?* had the same answer. Now, there are many different ways to answer those questions.

This OMB guidance encourages Federal agencies to collect aggregated information on a given population using the five single race categories and the four most common double race combinations. These four double race combinations are: (1) American Indian or Alaska Native and White, (2) Asian and White, (3) Black or African American and White, and (4) American Indian or Alaska Native and Black or African American. In addition to these categories, the March 2000 OMB guidance also encourages the aggregation of data on any multiple race combinations that comprise more than one percent of the population of interest to the Federal agency. OMB's guidance also encourages the reporting of all remaining multiple race data by including a "balance'' category so that all data sum to 100 percent.

In March 2000 the Office of Management and Budget (OMB) issued "Guidance on Aggregation and Allocation of Data on Race for Use in Civil Rights Monitoring and Enforcement" (OMB Bulletin No. 00-02, March 9, 2000; http://www.whitehouse.gov/omb/bulletins/b00-02.html). This document presented guidance on how to report race/ethnicity data from the 2000 Census.

According to the OMB, "[t]his method keeps intact the five single race categories, and includes the four double race combinations most frequently reported in recent studies. The method also provides for the collection of information on any multiple race combinations that comprise more than one percent of the population of interest." The following is an *example* that illustrates this guidance:

OMB Office of Civil Rights Reporting	
1	American Indian or Alaska Native
2	Asian
3	Black or African American
4	Native Hawaiian or Other Pacific Islander
5	White
6	American Indian or Alaska Native *and* White
7	Asian *and* White
8	Black or African American *and* White
9	American Indian or Alaska Native *and* Black or African American
10	> 1 percent: Fill in if applicable[1]_____
11	> 1 percent: Fill in if applicable_____
12	Balance of individuals reporting more than one race
13	Total

While this guidance presents a methodology for reporting **racial** data, Hispanic can easily be added so that all data are reported.

"Federal agencies will use the following rules to allocate multiple race responses for use in civil rights monitoring and enforcement.

.   Responses in the five single race categories are not allocated.
.   Responses that combine one minority race and white are allocated to the minority race.
.   Responses that include two or more minority races are allocated as follows:
    .   If the enforcement action is in response to a complaint, allocate to the race that the complainant alleges the discrimination was based on.
    .   If the enforcement action requires assessing disparate impact or discriminatory patterns, analyze the patterns based on alternative allocations to each of the minority groups.
.

Allocation for enforcement purposes should not be confused with various allocation methods under consideration for "bridging" to past data collections as described in OMB's Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity. These bridging methods would take advantage of information being gleaned from Census 2000 and other experimental work being carried out by the statistical agencies. The principal purpose of allocation for bridging is to conduct trend or time series analysis."

TAIR DAC/Ruddock

)

While bridging methodologies are used for presenting trend data and "allocation", in the federal terms refers to data displays, an institution may wish to use a bridging data display. There are two ways to present data using the "all inclusive" or duplicative methodology for allocating responses to a single race category – solely all inclusive, or all inclusive plus excluding single responses.

All Inclusive (Duplicated Headcount) by Race	
White/White + Any Combination of Race	White alone, and any combination of White with Black, Hispanic, Asian, American Indian, and Native Hawaiian
Black/Black + Any Combination of Race	Black alone, and any combination of Black with White, Hispanic, Asian, American Indian, and Native Hawaiian
Asian/Asian + Any Combination of Race	Asian alone, and any combination of Asian with White, Black, Hispanic, American Indian, and Native Hawaiian
Nat Hawaiian/Nat Hawaiian + Any Combination of Race	Native Hawaiian alone, and any combination of Native Hawaiian with White, Black, Hispanic, Asian, and American Indian
Am Indian/Am Indian + Any Combination of Race	American Indian alone, and any combination of American Indian with White, Black, Hispanic, Asian, and Native Hawaiian
Hispanic	Hispanic alone, and any combination of Hispanic with White, Black Asian, American Indian, and Native Hawaiian
Total - all inclusive by race	*The total number would be greater than the total number of respondents, because each respondent may be counted in one or more categories*
Hispanic	
Total with Hispanic	

This data display answers the question "how many of your students consider themselves to be _____?". The total number here (although you may choose not to add a total), would be greater than the number of respondents.

All Inclusive (Duplicated Headcount)	
White/White + Any Combination	White alone, and any combination of White with Black, Hispanic, Asian, American Indian, and Native Hawaiian
White and other races/ethnicity	Any combination of White with Black, Hispanic, Asian, American Indian, and Native Hawaiian
Black/Black + Any Combination	Black alone, and any combination of Black with White, Hispanic, Asian, American Indian, and Native Hawaiian
Black and other races/ethnicity	Any combination of Black with White, Hispanic, Asian, American Indian, and Native Hawaiian
Asian/Asian + Any Combination	Asian alone, and any combination of Asian with White, Black, Hispanic, American Indian, and Native Hawaiian
Asian and other	Any combination of Asian with White, Black, Hispanic,

races/ethnicity	American Indian, and Native Hawaiian
Nat Hawaiian/Native Hawaiian + Any Combination	Native Hawaiian alone, and any combination of Native Hawaiian with White, Black, Hispanic, Asian, and American Indian
Native Hawaiian and other races/ethnicity	Any combination of Native Hawaiian with White, Black, Hispanic, Asian, and American Indian
Am Indian/Am Indian + Any Combination	American Indian alone, and any combination of American Indian with White, Black, Hispanic, Asian, and Native Hawaiian
American Indian and other races/ethnicity	Any combination of American Indian with White, Black, Hispanic, Asian, and Native Hawaiian
Hispanic/Hispanic + Any Combination	Hispanic alone, and any combination of Hispanic with White, Black Asian, American Indian, and Native Hawaiian
Hispanic and any race	Any combination of Hispanic with White, Black Asian, American Indian, and Native Hawaiian
Total - all inclusive (including ethnicity)	*The total number would be greater than the total number of respondents, because each respondent may be counted in one or more categories*
Total - excluding single responses	

This data display adds information on how many respondents were in the single-race categories, and how many were in multiple-race categories. The total number here (although you may choose not to add a total), would be greater than the number of respondents.

The Deterministic Fractional Assignment (Equal Fractions) bridging methodology would produce a data display that looks like this:

Deterministic Fractional Assignment (Equal Fractions)			
		Fractional assignment	percent
White/White + Any Combination of Race	White alone, and any combination of White with Black, Hispanic, Asian, American Indian, and Native Hawaiian		
Black/Black + Any Combination of Race	Black alone, and any combination of Black with White, Hispanic, Asian, American Indian, and Native Hawaiian		
Asian/Asian + Any Combination of Race	Asian alone, and any combination of Asian with White, Black, Hispanic, American Indian, and Native Hawaiian		
Nat Hawaiian/Nat Hawaiian + Any Combination of Race	Native Hawaiian alone, and any combination of Native Hawaiian with White, Black, Hispanic, Asian, and		

	American Indian		
Am Indian/Am Indian + Any Combination of Race	American Indian alone, and any combination of American Indian with White, Black, Hispanic, Asian, and Native Hawaiian		
Hispanic	Hispanic alone, and any combination of Hispanic with White, Black Asian, American Indian, and Native Hawaiian		
		Adds to number of respondents	*Adds to 100%*

Because fractional assignment was used here, the numbers for fractional assignment would add up to the total number of respondents because every respondent is being assigned to only one category based on an equal fractional weighting.

The federal government has also issued guidelines on such things as what names and abbreviations to use, as well as how to handle small cell sizes. This information can be found in the NCES Statistical Standards report issued August 1, 2008 (http://nces.ed.gov/statprog/2002/std1_5.asp).

Small Cell Sizes
"When reporting data on race and ethnicity in government publications, every effort must be made to use at least the minimal reporting categories, … whenever possible. More categories should be used when there are enough cases to support finer detail. However, if there are not enough cases in any individual category of race or Hispanic ethnicity, the data for that category and for the next smallest category must be included in the total but not shown separately, and must be footnoted as such. Alternatively, if several categories cannot be shown, the combined categories must be reported as an "other" category, and footnoted to describe the exact components." STANDARD 1-5-4

"In cases where there are too few cases to report all race/ethnicity groups separately or where there is a justifiable substantive reason to aggregate categories the aggregated category may be labeled using one of the following options:

List the names of the categories included in the aggregation (e.g., Black, Hispanic, Asian, Pacific Islander, American Indian);List the name of one of the categories followed by a reference to 'Other Races/Ethnicities' (e.g.," Black or African American and Other Races/Ethnicities"); orUses label 'All other Races/Ethnicities'" STANDARD 1-5-4

Presenting Data
"When using aggregations of race/ethnicity groups, the terms 'nonwhite' and 'other than White' are **not** acceptable for use in the presentation of Federal Government data. Similarly, do not identify or designate aggregations of race/ethnicity groups as 'minority' groups." STANDARD 1-5-4

"The names for the groups should be capitalized, per the U.S. Government Printing Office (e.g., White, Black, Asian, etc.)." GUIDELINE 1-5-4A

"The following abbreviated names are suggested for use in text or in tables and figures:American Indian (instead of American Indian or Alaska Native), Black (instead of Black or African American), Pacific Islander (instead of Native Hawaiian or Other Pacific Islander), Hispanic (instead of Hispanic or Latino).

A footnote is needed to describe these 'abbreviations' as follows:American Indian includes Alaska Native, Black includes African American, Pacific Islander includes Native Hawaiian, and Hispanic includes Latino. Race categories exclude Hispanic origin unless specified".GUIDELINE 1-5-4B

Appendix A
Office of Management and Budget Bridging Methodologies

Deterministic Whole and Fractional Assignment
In this example, there are three whole assignment tables (based on different assignment criteria), and one fractional assignment (Table C). Deterministic whole assignments are methods that use fixed, deterministic rules for assigning multiple responses back to one and only one of the racial categories from the previous standards.

The three types of **Deterministic Whole Assignment** presented here (taken from the OMB report) are:

1. The **Smallest Group** assigns responses that include White and another group to the other group, but responses with two or more racial groups other than White are assigned into the group with the fewest number of individuals identifying that group as a single race.
2. The **Largest Group Other Than White** assigns responses that include White with some other racial group, to the other group, but with responses with two or more racial groups other than White as assigned into the group with the highest single-race count.
3. The **Largest Group** assigns responses with two or more racial groups into the group with the largest number of individuals as a single race.
4. The **Plurality** alternative assigns responses based on data from the National Health Interview Survey.

The **Deterministic Fractional Assignment** is a method using fixed, deterministic rules for fractional weighting of multiple-race responses; that is, assigning a fraction to each of the one of the individual race categories that are identified. In this analysis, the **Deterministic Equal Fractions** methodology was used (Table D). Each of the multiple responses are assigned in equal fractions to each racial group identified. Thus, responses with two racial groups are assigned half to each group; those with three groups are assigned one-third to each; etc. The advantage of this methodology is that it is easily described. On the negative side, it does not take into account the current distribution of race and ethnicity that is not, in the case of most institutions, equally proportionate. Using this methodology, it is also impossible to link race/ethnicity data to other variables such as birth rate. This is not as much an issue for higher education as it is for health services or other agencies.

The **Deterministic Weighted Fraction** methodology uses previously determined fractions for assigning responses to each racial group identified. In the OMB study the data for weighting was taken from the National Health Interview Survey, where "respondents were permitted to select more than one race with only the first two responses captured. NHIS respondents reporting more than one race were given a follow-up question asking them for the one race with which they most closely identify. …. All responses in a particular multiple-race category using the Plurality

method are assigned to the race group with the highest proportion of responses on the follow-up question about main race."   No data are presented here. Discussion would be needed on what proportions to use.

Probabilistic Whole and Fractional Assignment
According to the OMB, **Probabilistic Whole Assignment** uses "probabilistic rules for assigning multiple race categories back to one and only one of the previous racial categories. Two alternatives are examined. They parallel the two alternatives discussed under Deterministic Fractional Assignment, except that, for a given set of fractions, the response is assigned to only one racial category. The fractions specify the probabilities used to select a particular category. The first alternative uses equal selection probabilities. The second uses the NHIS fractions where possible, and equal fractions when no information is available from NHIS. *Probabilistic Whole Assignment will yield nearly, on average, the same population counts as Deterministic Fractional Assignment.*"

According to the OMB, the **Probabilistic Fractional Assignment** method is not used because the alternatives are unnecessarily complex and do not improve upon the alternatives in the other cells (methodologies).

The hot deck imputation method is often used in surveys to provide data on responses to survey items where responses are missing. "For purposes of bridging, a hot deck procedure would find the 'nearest neighbor' on a number of demographic dimensions for a person who identified more than one racial group. The person would then be assigned into one of the racial categories that he or she had reported based on the single racial group reported by the nearest neighbor." This methodology was not given further discussion in the OMB report.

Appendix B
Texas State Demographer's Bridging Methodology

The Office of the State Demographer for Texas has prepared a paper on "Comparing Race/Ethnicity Between the 2000 Census and Earlier Censuses" (http://txsdc.utsa.edu/txdata/redistrict/re-report.php). This paper presents assumptions used to bridge data from the 2000 Census (the "new" categories) back to the 1990 census categories (the "old/current" categories). However, concerning this methodology, it must be noted that the 2000 Census allowed respondents to fill in "some other race", an option not allowed for the IPEDS or Department of Labor reporting.

Previous work on race and ethnicity conducted by the Texas State Data Center and the Office of the State Demographer have concentrated on four large mutually exclusive race/ethnicity groups – Anglos (non-Hispanic whites), Blacks (non-Hispanic Blacks), Hispanics (of all races), and an Other category which consists of non-Hispanic persons from all other racial categories, with a majority of this group consisting of Asians. For the 2000 Census data, they continue to look at these categories.

These assumptions are:
1. Assumptions About the Hispanic Population
    It is assumed that all persons of Hispanic origin of all races should be considered Hispanic for purposes of comparison. The rationale is that for 1990, "the three categories of White and Some Other Race, alone and in combination with each other, account for 97.7 percent of all Hispanics. Only 0.7 percent indicate they are American Indian or Alaska Natives, 0.6 percent indicate that they are Black, 0.1 percent Asian, and 0.06 percent indicate that they are Native Hawaiian and Pacific Islander. That is a total of 99.2 percent of all Hispanics are in these seven categories. Only 0.8 percent are in the remaining 56 categories."
2. Assumptions About Non-Hispanic Populations Indicting Membership in Three or More Race Groups
    "Forty-two of the 63 racial groups involve persons who identify themselves as members of 3 or more racial/ethnic groups. However, in Texas only 0.093 percent of the non-Hispanic population (and, as noted above, only 0.09 percent of the total population) was in such groups. These groups are very diverse and cannot be easily allocated to any of the Hispanic, Anglo or Black groups…. Given the small size and diversity of these groups, we allocate persons in all of these three or more racial group categories of non-Hispanics to the Other category."
3. Assumptions About Single Race Groups of Non-Hispanics
    "Consistent with 1990 we allocate these single race non-Hispanics to the four racial/ethnic groups we have used historically as follows: non-Hispanic Whites are allocated to the Anglo; Non-Hispanic Blacks to the Black category;

and Non-Hispanic persons of American Indian and Alaska Native, Asian, Native Hawaiian and Other Pacific Islander, or Some Other Race to the Other category."
4. Assumptions About Non-Hispanic Populations Indicting Membership in Two Race Groups Only

Each of the four alternatives presented contain the same assumptions for handling the Hispanic population, non-Hispanic populations indicating membership in three or more race groups, and single race groups of non-Hispanics.

Alternative I

non-Hispanics of all other single and multiple racial groups alone or in combination can be categorized as Other.

Alternative II

Persons in the category of non-Hispanic White and Some Other Race were considered to be Anglo; persons who are non-Hispanic Black and Some Other Race are categorized as Black; and non-Hispanic persons in the American Indian or Alaska Native, Asian, Native Hawaiian and Pacific Islander, and Some Other Race were placed in the Other racial/ethnic group category. All other two and three race combinations of non-Hispanics are also placed in the Other non-Hispanic category

Alternative III

a. Persons in the category of non-Hispanic White and Some Other Race were considered to be Anglo; persons who are non-Hispanic Black and Some Other Race are categorized as Black; and non-Hispanic persons in the American Indian or Alaska Native, Asian, Native Hawaiian and Pacific Islander, and Some Other Race were placed in the Other racial/ethnic group category. All other two and three race combinations of non-Hispanics are also placed in the Other non-Hispanic category, with the exception in e.
b. The two-race combination of White/Black is allocated from the Other category to the Black category

Alternative IV

a. The two-race combination of White/Black is allocated to the Black category; White and any other stand alone category is allocated to the White category; Black and any other stand alone category is allocated to the Black category
b. All other combinations are allocated to the Other category

The State methodology relies on a sequencing of choices for allocating/bridging respondents to different categories, with Alternative IV representing the most comprehensive allocation.

Appendix C
Bridging Methodologies (Consortium on Financing Higher Education)

Broh and Minicucci(2008) looked at the fifteen two-category responses and constructed bridging strategies for those responses. Their rules are:

10. The classification applies only to US Citizens and Permanent Residents.
11. Respondents who answer the race/ethnicity question, but not the citizenship question are assumed to be US citizens or Permanent Residents; foreign includes those who responded to the citizenship question as "Foreign" (this applies to the Common Application for private institutions and reporting to the College Board)
12. Those not responding to either the first or second question are summarized as "Unknown"
13. Hawaiian and other Pacific Islander are classified as "Asian"
14. Those selecting more than two categories are classified as "more than two"
15. Black supersedes other identities; that is, a respondent choosing African American and *any other combination* of identities is summarized as Black
16. Hispanic supersedes identities other than Black
17. Asian supersedes White
18. American Indian identities do not supersede any other category

Like the Texas State Data Center report, the Broh and Minicucci(2008) article concentrates on the bridging from a two-race category to a one-race category. Based on research, in their methodology:

Latino-White (the largest combination of two race and ethnicity categories) is allocated to Latino

Asian-White (the second largest) is allocated to Asian because "anecdotally, Asian students are likely to be undercounted because of the belief that affirmative action policies discriminate against the Asian population."Counting these students as Asian preserves the information about the non-White, Asian identity of the student.

Black-White is allocated to Black because "preserving information about this group of underrepresented minorities to understand affirmative action policies is a high priority for a bridging strategy."

American Indian-White is allocated to White. Conceptually, American Indian is different from all other categories because of the requirement (on the Census and the Common Application) for the respondent to state their tribal affiliation.

Latino-Black is allocated to Black based on research conducted by COFHE that showed COFHE admissions officers counted these students as Black.

Latino-Asian is allocated to Latino to retain information on underrepresented students.

Asian-Hawaiian is allocated to Asian to follow the pre-2000 category that

> combined these two categories under the name Asian.
> Asian-Black is allocated to Black to retain information on underrepresented minorities.
> American Indian-Black is allocated to Blackto retain information on underrepresented minorities.
> Latino-American Indian is allocated to Latino.
> Latino-Hawaiian is allocated to Latino.
> American Indian-Asian is allocated to Asian.
> Black-Hawaiian is allocated to Black for recruitment and affirmative action purposes.
> American Indian-Hawaiian is allocated to Asian.

According to Broh and Minicucci (2008), "COFHE's assignment of students with two racial or ethnic identities to a single category relies on a set of 'trumping rules' that are different from those required in the Guidance [the federal reporting rules]. The IPEDS trumping rules have Latino trumping each of the other race categories; COFHE trumping rules reflect historical trends and conceptual information that has historically guided higher education policy for three decades."

"Racial Identity and Government Classification: A Better Solution" presented by Anthony Broh and Stephen Minicucci at the 2008 forum of the Association of Institutional Research.