

Racial Identity and Government Classification:  
A Better Solution

by

C. Anthony Broh  
Director of Research Policy  
Consortium on Financing Higher Education

Stephen D. Minicucci  
Director of Analytic Studies  
Consortium on Financing Higher Education

Paper presented at the  
48<sup>th</sup> Forum of the  
Association for Institutional Research

Seattle, Washington  
May 28, 2008

## **Racial Identity and Government Classification: A Better Solution**

C. Anthony Broh  
Director of Research Policy  
Consortium on Financing Higher Education

Stephen D. Minicucci  
Director of Analytic Studies  
Consortium on Financing Higher Education

### **ABSTRACT**

On October 19, 2007, the Office of Management and Budget issued its long awaited “Final Guidance on Maintaining, Collecting, and Reporting Racial and Ethnic Data to the U.S. Department of Education.” The “Guidance” imposes, for the first time, a required format for COLLECTING data on race. This format is flawed but it is a *fait accompli* to which institutions of higher education must adapt. This paper outlines current efforts from the College Board, Common Application, and the Consortium on Financing Higher Education to provide data collection instruments that comply with federal standards and meet the evolving needs of college admissions officers and administrators. The paper also describes data STORAGE requirements. Finally, new standards for Federal REPORTING of race and ethnicity data are also being imposed. This, too, is a flawed regulation which will result in considerable loss of information compared to previous practices, especially about Asians and African Americans. Here, we recommend institutional reporting practices that preserve, to the extent possible, both the maximum amount of information about the race and ethnicity of students and continuity with institutional trend data. Institutions should begin planning soon to prepare for these changes.

## **Racial Identity and Government Classification: A Better Solution**

by

C. Anthony Broh  
Director of Research Policy  
Consortium on Financing Higher Education

Stephen D. Minicucci  
Director of Analytic Studies  
Consortium on Financing Higher Education

The “Final Guidance on Race and Ethnicity<sup>1</sup>” imposes a flawed scheme for collecting data about student race and ethnicity and marks a significant break with current practices in higher education. It also establishes a reporting system which will result in the loss of significant amounts of information, especially about Asian and African American students. The Guidance is, however, a *fait accompli*; the issue for higher education now is how best to comply with the mandate with minimal information loss.

After underscoring the unprecedented nature of the collection requirements that the Guidance establishes, this paper describes the two central elements of the new mandates. It then describes some on-going efforts to design data collection processes that meet institutional needs while complying with the Guidance, outlines some considerations related to the storage of data on student race and ethnicity, and makes a set of recommendations regarding how institutions should report on race and ethnicity for their own purposes.

### **THE “GUIDANCE”**

The 2008 Guidance marks the first time the federal government imposed data collection requirements on institutions of higher education. In previous regulations such as Title IV of the Higher Education Act, colleges and universities had to report data about race and ethnicity to the Department of Education to remain eligible for various types of federal dollars,<sup>2</sup> but the techniques and methods for collecting the data were not specified. Institutions did, and continue to gather the information very differently from government practices and use procedures that respond to local administrative needs and processes.

---

<sup>1</sup> Department of Education, “Final Guidance, on Maintain, Collecting and Reporting Racial and Ethnic Data to the Department of Education,” *Federal Register*, Vol.72, No. 202, October 19, 2007. Henceforth, the “Guidance.”

<sup>2</sup> Indeed this requirement dates only to the 1992 reauthorization of the Higher Education Act. For many years the NCES refused to be an enforcement mechanism believing it could not have “objective data collection” if it had an enforcement responsibility. In general, Title IV violations result in fines rather than loss of all Title IV funding. Fines for non-completion of IPEDS were added in the 1998 reauthorization.

For the first time in 2008, the government specified the format and the timing for collection in addition to the recognized obligation to collect the data for reporting.

The original 1977 standards that specified how racial and ethnic data should be collected, known as “Office of Management and Budget Statistical Directive 15” (or “OMB 15” for short),<sup>3</sup> instructed the departments and agencies of government about the nomenclature and collection format. The collection strategies were also specified for surveys or data forms, but the regulations only related to government activities or federally funded projects. When the government collected data from third parties or organizations, OMB 15 was silent. Institutions were free to collect the data according to the principles that were the most useful for the policies they intended to observe and evaluate.

Indeed, the 1997 Guidance, a reissue of OMB 15 policy, and its 2001 interpretations continued the practice of not requiring institutions of higher education to collect data in a specified format. The 1997 Guidance, which updated the format for data collection in government agencies and contracts, remained silent about whether the collection format was a requirement for third parties. With the data experiments of Census 2000 as support, the Office of Management and Budget in the last few days of the Clinton administration published its interpretation of the 1997 Guidance for different segments of the economy. Again collection formats were not a requirement.

Of particular significance in the 2001 interpretations is the recognition that different segments of the economy might need to collect racial and ethnic data according to different objectives. OMB 15 allowed choice by specifying two different methodologies, even for government agencies and federally funded projects. For example, the so-called two-question format had become a common practice for collection in economic spheres, such as the Bureau of Labor Statistics and the Current Population Surveys of the U.S. Census Bureau. K-12 and post-secondary education, on the other hand, used the so-called one-question format for collecting data about race and ethnicity. The 2001 interpretations recognized this distinction, the burden for change for higher education, and encouragement for allowing the same choice as OMB 15 had allowed.

As recently as 2006, the federal government still had not imposed a collection requirement on higher education or other segments of the economy. For example, the guidance that governs personnel issues and federal government statistics about employment only “recommend” a data collection format.<sup>4</sup> It recognizes that data continuity and trend lines are important statistical indicators of compliance with federal policy, as well as possible litigation related to affirmative action.

When the Department of Education asked for comment on the proposed version of the Guidance in 2006, a collection format was still only recommended. In November 2006,

---

<sup>3</sup> Office of Management and Budget, “Directive No. 15: Race and Ethnic Standards for Federal Statistics and Administrative Reporting,” May 12, 1977.

<sup>4</sup> Equal Employment Opportunity Agency, “Agency Information Collection Activities: Notice of Submission for OMB Review, Final Comment Request,” *Federal Register*, Vol 70, No. 227 (November 28, 2005), pp. 71294-71303.

discussion and materials for a meeting of a Technical Review Panel that had been convened by the Department of Education continued to indicate that collection formats in the Final Guidance would be recommended. The authors of this paper commented officially to the Secretary that a collection format in the Guidance should remain “recommended” so that colleges and universities could collect data according to their local administrative needs and practices while remaining in compliance with the needs of the Department of Education for public reporting and Congressional oversight.

The fact of a new requirement on higher education that may or may not be sensitive to institutional needs generally happened without debate among the higher education research community. This paper documents that omission with an objective of minimizing the administrative damage for colleges and universities while also following the Guidance to ensure institutional compliance with Title IV requirements.

Over the decade-long discussion leading up to the Guidance, government employees from National Center for Education Statistics (NCES) assured the research community that it would allow time for implementing new standards. It self-imposed a three year transition period for colleges and universities to develop collection forms, administrative procedures and data systems that comply with the new guidelines. Reports of the Integrated Postsecondary Education Data System (IPEDS) have different lag periods for compliance depending upon the logical time frame in which the institutional data might be collected. Data from enrolled students, for example, can be collected more readily than data from graduating seniors who might or might not be in touch with administrators until the point of graduation. Thus IPEDS reports about fall enrollments are due to NCES a year prior to reports about degree completions with the earlier following a three year lag after publication of the Guidance.<sup>5</sup>

With an implementation period of three years, many colleges and universities thought about delaying work on implementation as long as possible. Indeed one author of this paper encouraged colleges and universities to do just that.<sup>6</sup>

However, institutional researchers should begin a planning process where three years is a short period. Specially, computer programmers and software vendors need to know the format requirements of the data so they can adapt legacy systems for storage and reporting requirements. Additionally state agencies and publishers require lead time to know how to collect the data from colleges and universities so they can adapt their own collection and reporting practices.

In sum, the Guidance for Data about Race and Ethnicity has several new components and imposes standards on institutions of higher education that previously were only recommended or not specified at all. While the tendency of many for a new obligation is

---

<sup>5</sup> See, Department of Education, “Supporting Statement for IPEDS 2007-2010, OMB Paperwork Reeducation Act Submission,” May 4, 2007.

<sup>6</sup> C. Anthony Broh, “Race to the Unknown,” *Transcript*, An online News Source for members of the American Association for Collegiate Registrars and Admission Officers  
[http://www.aacrao.org/transcript/index.cfm?fuseaction=show\\_view&doc\\_id=3503](http://www.aacrao.org/transcript/index.cfm?fuseaction=show_view&doc_id=3503)

to ignore or delay compliance, institutional researchers should begin preparing now for 2010 when the first reporting requirements become mandatory.

## TWO DATA COLLECTION ISSUES

The Guidance mandates two major changes in the way that institutions collect data about student race and ethnicity. The first is the use of the two-question format and the second is the use of the “mark one or more” instruction for selecting a race category. Presently, few institutions have data collection procedures that comply with both of these rules, so the new rules will require significant adjustments.

### **Issue One: The two-question collection format replaces the one-question collection format.**

The Guidance requires institutions to gather data about race and ethnicity with two survey questions rather than one question that includes all races and Latino. The first question of the two question format requests a “yes” or “no” response to “Are you Hispanic/Latino?” -- defined in the 1997 standards as follows:

Hispanic or Latino<sup>7</sup>. A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. The term, "Spanish origin," can be used in addition to "Hispanic or Latino."

The second question allows respondents to select among five racial categories: (1) American Indian or Alaska Native; (2) Asian American<sup>8</sup>; (3) Black<sup>9</sup> or African American; (4) Native Hawaiian or Other Pacific Islander; and (5) White. The official descriptions of each group follow:

American Indian or Alaska Native. A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

---

<sup>7</sup> The authors of this paper note that both the words “Hispanic” and Latino” are considered English language words. “Hispanic” is actually a term that was coined in the 1970s to describe a group that was originally identified by their “Spanish surname.” While “Latino” has its translation in the Spanish language and is masculine in gender, English language nouns do not have gender. English pronouns do have gender so it is not an irrelevant concept. Thus “Latino” in English language data collection refers to both males and females while Spanish language data collection should use the masculine (“Latino”) and feminine (“Latina”) nomenclature, such as “Latino/a.”

<sup>8</sup> The nomenclature for the racial and ethnic categories in this paper conforms to the proposed government guidance with two exceptions. The data section of this paper refers to a survey instrument that used slightly different nomenclature: “Asian American” rather than “Asian” to include the assumption of citizenship and “Hispanic/Spanish Origin or Latino/a” instead of “Hispanic or Latino.”

<sup>9</sup> The Guidance states that all racial and ethnic categories begin with an upper case letter. In this paper, “Black” and “White” refer to racial classifications while “black” and “white” refer to the peoples with these identities.

Asian. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

Black or African American. A person having origins in any of the black racial groups of Africa. Terms such as “Haitian” or “Negro” can be used in addition to “Black or African American.”

Native Hawaiian or Other Pacific Islander. A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

White. A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

Prior to the Guidance, institutions of higher education almost universally exercised their discretion within OMB Statistical Directive 15 (1977) and used one question to record race and ethnicity with “Hispanic” listed as an item among the other race categories.

While the main focus here is on how best to make the transition to the new requirements, it is worth underscoring briefly that the research supporting the use of the two-question format is simply incorrect. The Department of Education claims that the two-question format “results in more complete [sic] reporting of Latino ethnicity,” but it never really tested its case. The experimental research did not substantiate the claim that the two-question format was more accurate.<sup>10</sup> Indeed, secondary analysis of the 1995 CPS Supplement data cited in the Guidance found that the two-question format did not produce more accurate data than the one-question format.<sup>11</sup> The 2006 panel at the National Academy of Sciences on *Hispanics and the Future of America* provides evidence that the two-question format was confusing to the young Hispanic population.<sup>12</sup> At the same time, it distorts the true identity of “underrepresented minorities” by setting Latino “above” black or American Indian. The two-question format makes the unfortunate suggestion that groups are not treated equally in higher education policy.<sup>13</sup>

The “one-question format” has been preferred because it avoids these sorts of confusions while providing all respondents the same opportunity to identify as Latino and/or with any race. Further, all data storage and reporting requirements for the one-question format can be met with the same storage requirements as the two-question format; there is no

---

<sup>10</sup> For more discussion, see C. Anthony Broh, “Race Matters,” Presentation at the Annual Forum of the Association for Institutional Research, Kansas City, Missouri, June 4, 2007

<sup>11</sup> *Ibid.*

<sup>12</sup> Marta Tienda and Faith Mitchell [ed], *Hispanics and the Future of America*, Report of the Panel on Hispanics in the United States, Committee on Population, Division of Behavioral and Social Sciences and Education. Washington, DC: National Research Council, 2006.

<sup>13</sup> This section is an abbreviated description of a letter sent to the Department of Education during the open comment period following publication of the preliminary guidance. A copy is available from the authors upon request.

technological need for the change. Most important, requiring higher education to use the two-question format is a major departure from prior practices.

### **Issue Two: Allowing Respondents to “Mark One or More” races.**

The second major change in data collection procedures imposed by the Guidance creates fewer problems for many institutions of higher education than the two-question format. Beginning in 1997, many standard data collection forms, including the U.S. Census 2000 forms, allowed respondents to “mark one or more” race categories – sometimes abbreviated as MOOM. By adopting this standard for data collection, then, the Department of Education is bringing its procedures into conformity with broader federal practices. Unlike the shift to the two-question format, this change in collection format does not distort the data. Indeed, many schools had allowed respondents to select more than one racial category on admission applications even prior to the 1997 standards and surveys at the Consortium on Financing Higher Education (COFHE) have allowed MOOM since 1997. While most of the national media and higher education associations focused on the change from single race collection to MOOM, the many years of experience with collection and storage using the MOOM standard posed little problem for a portion of schools.

The Guidance, then, does not create a significant break with current practices by requiring the use of MOOM in data *collection*; however it also requires that respondents marking more than one category be *reported* as “two or more.” This dramatic and destructive shift departs substantially from prior practice at most institutions of higher education. We return to that issue in the section of this paper about “reporting.”

## **CURRENT DATA COLLECTION EFFORTS**

Despite the flaws pointed out above, the collection format required by the Guidance will become the standard for higher education. Compliance is a necessary condition for post-secondary institutions to receive financial aid dollars under Title IV of the Higher Education Act. Consequently, the time has come to leave the objections behind and use the institutional necessity to reformat applications and forms as an opportunity to collect better racial and ethnic data.

In particular, changes in affirmative action policies resulting from legal advice after the 2003 Supreme Court decision in the University of Michigan cases<sup>14</sup> and from state referenda in California, Washington, and Michigan have modified the needs for data and information. For example, more detailed data on ethnicity could also be used in a holistic admissions process. Adding details is consistent both with the legal standard for the use of “race as a factor in admissions” and with the Guidance itself. The only

---

<sup>14</sup> Gratz v. Bollinger, 539 U.S. 244 (2003); Grutter v. Bollinger, 539 U.S. 306 (2003).

requirement for collecting information about subcategories is that they must “roll up” into the federally the mandated reporting schema.

Admissions data are both legally and strategically important for institutional data collection. Legally, federal policies regarding race and ethnicity in admissions, as well as student employment and associated litigation establish data collection efforts about students. Strategically, the admissions process is the point of entry for the information about race and ethnicity. At selective colleges and universities with active outreach programs, data collection processes are carefully constructed and maintained for benchmarking and evaluating the impact of admission committee decisions. At open enrollment institutions, admissions may be the only opportunity for collecting and verifying the information.

Admissions data are also the probable source of information about race and ethnicity for all enrolled students after admissions. Enterprise data systems, for example, store information about race and ethnicity centrally so it can be accessed from a single data table. At institutions where the admission system and the student record system are independent, information about race and ethnicity are typically rolled over to the student record system at the time of enrollment. Even if institutions allow enrolled students to update demographic information during registration or at other times, race and ethnic data from the admission system are typically the default value for each student.

Recognizing the starting point for the data, the Consortium on Financing Higher Education began working on its own survey collection formats and with the administrative data collected by two organizations that are the source of information in college admissions for many private institutions. In November 2006, the COFHE research staff met with representatives of the Common Application and the College Board to discuss collection formats.

As described on its website, “the Common Application membership association was established in 1975 by 15 private colleges that wished to provide a common, standardized first-year application form for use at any member institution.... Now in [its] fourth decade, the Common Application currently provides both online and print versions of its First-year and Transfer Applications...for more than 300 institutions...in the US: public and private, large and small, highly selective and modestly selective, and East Coast, West Coast, and every region in between.”<sup>15</sup> Relevant to the discussion here, the Common Application collects data about race and ethnicity in its application and this information is typically downloaded into campus information systems.

“Founded in 1900, the [College Board] is composed of more than 5,400 schools, colleges, universities, and other educational organizations. Each year, the College Board serves seven million students and their parents, 23,000 high schools, and 3,500 colleges through major programs and services in college admissions, guidance, assessment, financial aid, enrollment, and teaching and learning.”<sup>16</sup> Of particular relevance to this

<sup>15</sup> <https://www.commonapp.org/CommonApp/History.aspx>

<sup>16</sup> <http://www.collegeboard.com/about/index.html>

discussion, the College Board administers the SAT Exam that many private institutions require for admissions. The online application to take the Exam collects information about race and ethnicity that participating institutions may also download into their data systems.

Colleges and universities that engage in active recruitment of underrepresented minorities can subscribe to a College Board service known as “College Search.” Using criteria such as a minimum SAT score, a geographic region of the country, and/or racial and ethnic identity, participating colleges and universities purchase mailing lists of students with specified demographic characteristics. For those schools that are seeking greater racial and ethnic diversity on campus, lists of underrepresented minority students typically populate a data base of prospective students whose data will then populate the admissions file if the student actually applies.

Like data from the Common Application, information from the College Board application becomes the source of data for many colleges and universities. Ensuring that the collection format conforms to the Guidance means that this information can be used for reporting to the Department of Education, as well as serving the needs of the admission office.

The collaboration among the Common Application, the College Board, and COFHE identified two objectives for data collection that previously were not necessarily synchronous in college admissions:

1. To meet the recruitment and admission objectives of the member institutions to each organization
2. To collect data that institutions could use for reporting to the Department of Education

A corollary to the first objective is to ensure that the manner will not offend applicants that are interested in applying to a school. Consequently, forms often have a provision for opting out of a question, even after an applicant may have begun answering it. For example, admission applications may have a category called “other” and admission officers sometimes advocated for a response such as “choose not to reply.” While “user friendly” for the person filling out an admission application or a request form, these responses do not comply with the Guidance and would therefore make the data ineligible for federal reporting. At this writing, the Common Application plans to employ focus groups to react to the proposed collection format described in this paper and the College Board is conducting experiments with collection formats; both are indicators of the admission office’s sensitivity to the college applicant population that institutional researchers may see unimportant or non-debatable.

Another major interest for admissions officers is a greater level of detail than the minimum requirements of the Guidance. This need suggests the inclusion of subcategories of race and ethnicity on admission forms, something that both the College Board and the Common Application are exploring. For example, the ancestry and

background of black students is an increasing focus of college admissions at selective colleges and universities. Campus research and public discourse about affirmative action suggest that second generation children of African immigrants are more likely to be admitted to some highly selective schools than African American applicants whose ancestry dates back to American slavery. At a recent Harvard University alumni event, Lani Guinier and Henry Louis Gates estimated that as many as a half to two-thirds of black students were immigrants of West Africa, the children of these immigrants, or the sons and daughters of biracial couples.<sup>17</sup> As Anthony W. Marx, president of Amherst College has observed,

... colleges should care about the ethnicity of black students because in overlooking those with predominantly American roots, colleges are missing an "opportunity to correct a past injustice" and depriving their campuses "of voices that are particular to being African-American, with all the historical disadvantages that that entails."<sup>18</sup>

Both objectives, affirmative action and federal reporting, are institutional responsibilities; however, the authority for each typically lies in different offices. Admission offices are responsible for the affirmative action objectives that include the collection of racial and ethnic data while institutional research offices are responsible for compliance with government reporting standards. Senior administrative officers at an institution may not recognize the differing objectives of the offices nor even the decentralized nature of the data collection and reporting responsibilities. Producing a consistent data collection format that meets both objectives serves the institutional interests while not compromising the responsibilities of each administrative office.

The Common Application developed the following format to meet both objectives:

**Common Application collection format for race and ethnicity**

1) Are you Hispanic or Latino?

- Yes, Hispanic or Latino (Including Spain)       No

*Which best describes your background?* ↓

Central America

Cuba

Mexico

Puerto Rico

South America (excluding Brazil)

Spain

Other \_\_\_\_\_

<sup>17</sup> Sara Rimer and Karen W. Arenson, "Top Colleges Take More Blacks, But Which Ones?" *The New York Times*, Education Section, April 21, 2008.

<sup>18</sup> *Ibid.*

2) Regardless of your answer to the prior question, please check *✓ one or more* of the following groups in which you consider yourself to be a member:

- American Indian or Alaska Native (including all Original Peoples of the Americas)

*Which best describes your background?* ↓

Alaska Native  
 Chippewa  
 Choctaw  
 Cherokee  
 Navajo  
 Sioux  
 Other \_\_\_\_\_

[Are you Registered?

- No
- Yes, please enter Registration number \_\_\_\_\_]<sup>19</sup>

- Asian (including Indian subcontinent and Philippines)

*Which best describes your background?* ↓

China  
 India  
 Japan  
 Korea  
 Pakistan  
 Philippines  
 Vietnam  
 Other East Asian \_\_\_\_\_  
 Other Indian Subcontinent \_\_\_\_\_  
 Other Southeast Asian \_\_\_\_\_

- Black or African American (including Africa and Caribbean)

*Which best describes your background?* ↓

African American  
 African  
 Caribbean  
 Other \_\_\_\_\_

- Native Hawaiian or Other Pacific Islander (Original Peoples)

*Which best describes your background?* ↓

Guam  
 Hawaii  
 Samoa  
 Other Pacific Islands (excluding Philippines) \_\_\_\_\_

- White (including Middle Eastern)

*Which best describes your background?* ↓

Europe  
 Middle East  
 Other \_\_\_\_\_

<sup>19</sup> This data element is part of the Common Application and emphasizes the requirement of “tribal affiliation” as a necessary component of American Indian identity in college admissions.

The collaboration of the College Board, the Common Application, and COFHE devised a set of principles that determined various decision rules for this collection format.

- The layout of the items minimizes the difference between the first and second question of the two question format.
- Race categories are listed in alphabetical order using nomenclature required in the Guidance.
- Explanatory material for racial and Latino identity is displayed as parenthetical expressions without altering the required nomenclature.
- In general, parenthetical material aids in making the categories mutually exclusive and provides detail for standards and definitions in the Guidance.<sup>20</sup>
- Subcategories collected on the web are displayed as drop-down menus with an instruction that avoids confusion between the meaning of race and ethnicity.
- Subcategories roll up into the racial and ethnic classification that the Guidance requires.
- Subcategories respond to admission officers' responsibility for racial and ethnic diversity.
- Subcategories reflect the largest United States populations in each racial and Latino identity.
- Where possible, the nomenclature for subcategories reflects geographic area rather than ethnic groups or religions (American Indian includes the names of the largest tribes).
- All subcategories include a residual "other" classification that allows additional information beyond that specifically listed, an opportunity to add information in free form, and avoidance of "other" as a major category in the classification.

If approved, the racial and ethnic classification for the Common Application will be first used with the entering class of 2009<sup>21</sup>.

The College Board approach to a collection format has a slower pace and is more research oriented. On February 21, 2008, SAT registrants were required to complete two race and ethnicity items during the online registration. The first uses the format that the College Board has used for several years and is not relevant to this discussion. The second item is a split-half experiment using the one-question format with nomenclature and categories (but not format) that are consistent with the Guidance and a two-question format with nomenclature and categories that are consistent with the Guidance. The two forms of the experiment (called "Spiral A" and "Spiral B" by the College Board) are reproduced below:

---

<sup>20</sup> One exception is the exclusion of immigrants of Portuguese speaking nations from the "Hispanic or Latino," which is explicitly defined as "Spanish speaking."

<sup>21</sup> At this writing, the Executive Committee of the Common Application has not met to give final approval. Modifications are possible by the time of the AIR Forum.

### One-Question Format in College Board 2008 Experiment

(SPIRAL A) 1. Race/Ethnicity – “Please check one or more of the following options that you identify with”

- American Indian or Alaska Native
- Asian or Asian American
- Native Hawaiian or Other Pacific Islander
- Black or African American
- Mexican or Mexican American
- Puerto Rican
- Other Hispanic, Latino, or Latin American
- White
- Other

### Two-Question Format in College Board 2008 Experiment

(SPIRAL B) 1. Are you Hispanic or Latino (including Spanish and other Spanish origin)?

- Yes, Mexican or Mexican American
- Yes, Puerto Rican
- Yes, Other Hispanic, Latino, Latin American
- No

(SPIRAL B) 2. Race/Ethnicity – “Please check one or more of the following options that you identify with”

- American Indian or Alaska Native
- Asian or Asian American (including Indian Subcontinent)
- Black or African American (including African and Afro-Caribbean)
- Native Hawaiian or other Pacific Islander
- White (including Portuguese, Brazilian, Persian, and Middle Eastern)

The College Board devotes less space on its application and registration procedures to the collection of race and ethnicity than the Common Application. Thus the number of subcategories has been compressed, but the nomenclature and parenthetical explanations generally conform to the decision rules used by the Common Application.<sup>22</sup>

Finally, the inclusion of subcategories in the collection format is important for an understanding of race and ethnicity. Currently, the Guidance recognizes only one ethnicity: Hispanic or Latino. All other categories in the classification are considered a race. Yet the defining characteristics of ethnicity typically include factors such as nationality, religion, or language. The Guidance implicitly recognizes at least one of these characteristics by separating “Spanish-speaking” groups from other data collection.

Yet the Guidance does not require data collection about ethnic differences among Asian American groups such as Chinese, Japanese, or Vietnamese, although it recognizes them

---

<sup>22</sup> An advisory committee of the Common Application made a few changes after the College Board experiment was in the field. Some nomenclature, such as “Persian,” was rejected by the Common Application, but Middle East remained on both collection formats.

in its definitions and explanations. Nor does it delineate important differences between the black population whose parents may have recently immigrated to the United States and the black population whose ancestors descended from generations of slavery in the United States. While one need not make decisions in the admission process based upon these distinctions, understanding ethnicity as part of a person's social background and identity is as important, perhaps more important, than the larger categories of race and ethnicity contained in the Guidance.

## STORAGE

The Guidance does not specify technical requirements about data storage in computer systems, but it identifies several tasks that might be necessary for future retrieval and reporting. In general, the regulations require the institution to preserve the original data that each student submitted. They acknowledge that a student identifying as African American and White might fall into one classification for government reporting on IPEDS but a different category in an EEOC complaint, a civil rights suit, or a federal grants program. Therefore, the institution needs to store the data in a manner that allows alternative calculations.

The practical implication for this requirement is the storage of each category of race and Spanish origin as separate data elements that record African American or not; American Indian or not, Latino or not, White or not, etc. Most, but not all, commercial administrative data systems are preparing for the Guidance and are likely to have this flexibility in new releases.

Although each data element requires only two values to record whether or not a student identifies with a race or ethnicity, the addition of subcategories (discussed above) suggests that each element might have multiple values that record the information with greater specificity. For example, a data element that records whether an applicant answered "yes" to the Latino question might have additional codes for Central American, Cuban, or Puerto Rican. IPEDS reporting, as well as specialized requests, legal requirements, or institutional reports requiring information that is not revealed from IPEDS are easily produced and maintained through a combination of data elements with multiple values.

In this sense, the data storage for the first question in the two-question format is the same as data storage for individual race categories in the second question of the two question format. That is, each race or ethnicity in the mandated collection requires a separate data element that stores information about the fact of identifying as Latino with various subcategories and the identification of each race with various subcategories. A key feature of any storage scheme is that it can be adapted to multiple reporting demands. The Guidance makes clear that IPEDS reporting to the Department of Education is not the only requirement for colleges and universities. Equal Employment Opportunity complaints, litigation, or so-called "bridging strategies" (discussed below) mean that information contained in the original collection categories, vis-à-vis information

contained in the Department of Education IPEDS categories, are required by campus administrators.

A storage system that places all people who responded to multiple race categories into “two or more” would not allow retrieval of data about the African American students who identified with more than one race; they would be mixed with other multi-racial students. Similarly, a storing data about Hispanic students without any indication of their racial categories, would not allow retrieval of data about African American students of Spanish origin. In each of these two scenarios, data storage about reporting categories without knowledge of the collection categories ignores additional combinations of race and ethnicity that are necessary for a variety of legal and administrative purposes.

Records with original information about a student’s racial and ethnic identity must be maintained for three years. This regulation is not a change in record keeping requirements; most schools already store original documents about admission, enrollment, and financial aid for this period and routine computer backup should also serve this purpose.

The Guidance did not require (but strongly recommended) so-called “resurveying,” in which each student enrolled at an institution is to be given the option of identifying with a race and ethnicity using the new collection format. Presumably, the racial and ethnic classification of existing students could default to the stored information about the student prior to the Guidance. For example, the data of, say, sophomores, juniors and seniors who used an old format that did not allow MOOM can be co-mingled with new data from entering first-year students and sophomores, juniors, and seniors after MOOM was required. Institutions are not required to store information from both formats.

Nevertheless, during the next three years of transition, the National Center for Education Statistics plans to allow submission of data from pre-Guidance or data from post-Guidance. Rolling over data from a pre-Guidance collection and storage into the post-Guidance categories would result in zero cases for Native Hawaiian and Other Pacific Islander since this category did not exist prior to the final Guidance. It would also result in “Other Pacific Islanders” in the pre-Guidance data becoming Asian in the post-Guidance data – a minor statistical adjustment for most institutions.

## REPORTING: THE “TWO-OR-MORE” PROBLEM

The most troubling element of the Guidance is the new requirement that respondents marking more than one “race” be reported in the category of “two-or-more.” The new convention marks a significant departure from current practices in higher education and will result in significant loss of information. Also, since the Department of Education rules for reporting Latinos depart from the rules for reporting racial categories, the Guidance creates an asymmetry among Latinos and other under-represented minorities.

The Guidance describes seven required reporting categories for IPEDS:

- (1) Hispanic/Latino of any race; and, for individuals who are non-Hispanic/Latino only
- (2) American Indian or Alaska Native
- (3) Asian
- (4) Black or African American
- (5) Native Hawaiian or Other Pacific Islander
- (6) White
- (7) Two or more races

Since these seven categories only apply to citizens and permanent residents of the United States, a category of “Foreign” students is also required to account for all students. The Guidance also allows continuation of an “unknown” category in IPEDS for students whose race and ethnicity can not be determined.

The problematic reporting category here is “two or more.” Table 1 identifies how the Guidance differs from current practices at Consortium institutions for cases where the respondent marked exactly two responses for race/ethnicity.

**Table 1: Selected Combination of Collection Categories with Reporting Requirements and Common Practices**

<b>COLLECTION response</b>	<b>GUIDANCE requirement</b>	<b>COMMON REPORTING practices<sup>23</sup></b>
Black and Latino	Latino	Black
Black and White	Two or more	Black
Black and Asian	Two or more	Black
American Indian and Latino	Latino	American Indian
Native Hawaiian and Asian	Two or more	Asian

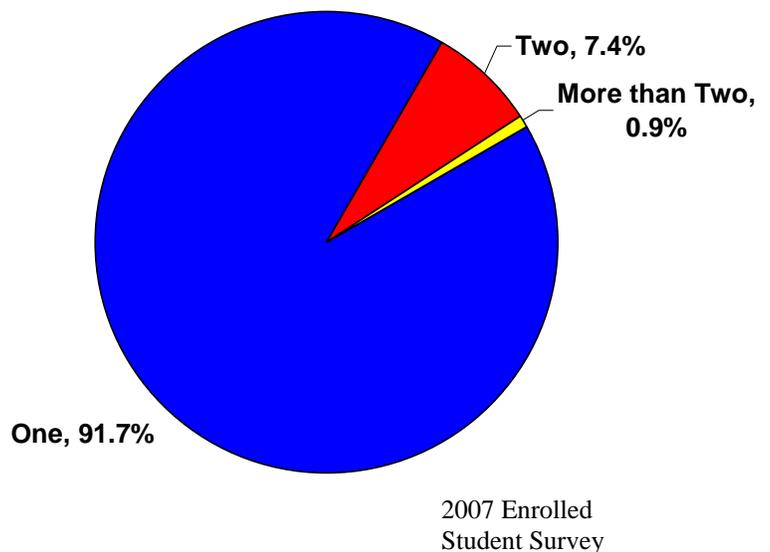
<sup>23</sup> C. Anthony Broh and Stephen D Minicucci, “One Organizations Bridging Strategy for Racial and Ethnic Classification,” Paper presented at the Annual Forum of the Association for Institutional Research, Tampa, FL, May 21, 2003.

It is easy to see in Table 1 that the counts for Latino students in IPEDS are unaffected by marking any other racial identity; Latino students are always Latino, regardless of their race. Institutional researchers often refer to decision rules like this—in which one classification has dominance over another classification—as “trumping.” In IPEDS reporting, Latino trumps all other races or ethnicities. The most contentious of these pairings is likely to be Latino-African American. Institutions that recruit these students as African American typically report them as African American. They now must be reported as Latino.<sup>24</sup>

When the pairing does not include the category “Latino,” a different reporting IPEDS rule applies. A student identifying as both White and African American might have been classified as African American for recruitment purposes. That student must, however, be counted in IPEDS reports as “two or more races.” The same is true for students indicating both American Indian and white. It is clear that the “two or more” convention will result in the loss of information about under-represented minorities.

To estimate the loss of information from IPEDS reporting more accurately, we tallied student reports from the 2007 COFHE Enrolled Student Survey. Figure 1 summarizes the number of student responses to the (“one-question” format) race/ethnicity question in the survey. Roughly one-in-twelve students (8.3%) of students at these private, highly selective colleges and universities identify with more than one race or ethnicity. In most of these cases, students marked exactly two responses. Less than one percent of all students chose more than two. In order to focus on the critical question of how the Guidance produces information loss, the remainder of this paper will set aside this last slice and focus on the 7.4% that chose exactly two races or ethnicities.

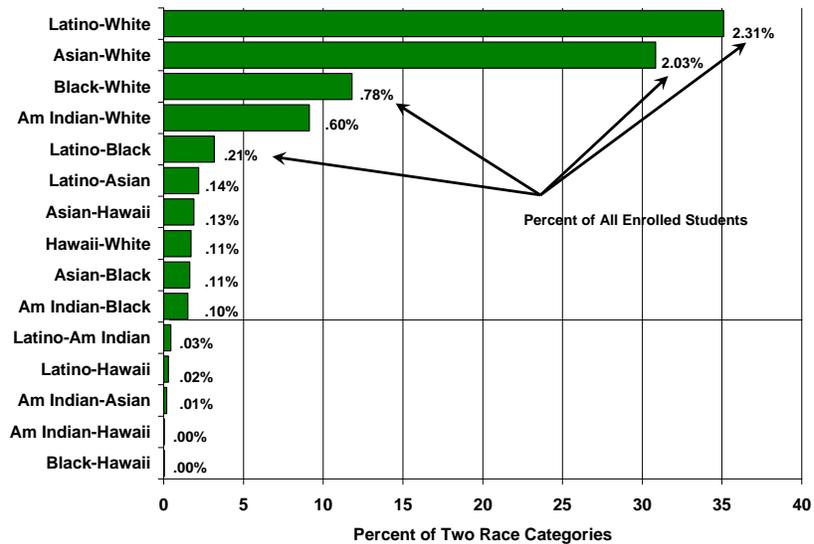
**Figure 1: Percentage of Respondents Marking One, Two, or More than Two Racial and Ethnic Categories.**



<sup>24</sup> Note that the concept of “trumping” is introduced in the Guidance itself, with Latino trumping all other categories in IPEDS reporting. However, the Guidance recognizes that other “trumping” strategies might be used for different purposes. We return to that discussion later in this paper.

Figure 2 displays the fifteen two-race and ethnicity combinations that result from the collection categories and their frequency from the 2007 Enrolled Student Survey. The percentages at the end of each bar in the graph are the percent of enrolled students in the category—this is an estimate of the loss of information that will result from implementation of the Federal reporting requirement to collapse all fifteen of these categories into a single group called “two or more.”

**Figure 2: Combinations of Responses Marking Two Races or Ethnicities**



For example, Figure 2 shows that COFHE schools will for federal purposes report an estimated 2.5% fewer Asians than prior years simply because of the requirement to count all dyadic combinations of Asian students as “two or more.” Similarly, it will report 1.2% fewer black students than prior years.<sup>25</sup>

### PREVENTING INFORMATION LOSS: “BRIDGING”

While the new IPEDS reporting guidelines will inevitably result in information loss, institutions need not lose this information in their *own* reporting. The Department of Education standards explicitly recognize that the new reporting requirement create a discontinuity with trend data collected and reported prior to the Guidance, and encourage institutions to use so-called “bridging strategies” of their own in calculating trend data for administrative, legal, and/or public reporting. Moreover, the Guidance allows institutions to collect additional information that might inform reporting categories for institutional purposes. The earlier section on collection suggested procedures that COFHE, the Common Application, and the College Board are using to meet the additional information needs of the higher education community.

Bridging strategies can provide a foundation for standardized reporting among schools that wish to produce trend lines that are responsive to institutional objectives and overcome the information loss inherent in the IPEDS reporting requirements. The proposed standards from August 2006 even offered examples that are directly relevant to the discussion here:

<sup>25</sup> These data are exemplary of what happens to COFHE’s 31 member institutions based upon survey results. COFHE does not report data about itself or member institutions for IPEDS.

For example, multiple race responses that combine one minority race and White could be allocated to the minority race.<sup>26</sup>

The (final) Guidance in October 2007 offers several suggestions for bridging strategies and provides examples of situations where schools may wish to make alternative calculations to the IPEDS standards:

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 categories and relate them back to the racial 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.<sup>27</sup>

In 2003, the authors of this paper created a summary of the desirable characteristics for a bridging strategy, which we applied to a variety of alternatives referenced in the 1997 Standards.<sup>28</sup> They are relevant to a bridging strategy that complies with the Final Guidance:

- 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.

These same characteristics are important criteria for deciding on a bridging strategy that recovers the lost information in the IPEDS reporting format. Figure 2 displayed fifteen combinations where racial and ethnic information is lost through the format in the Guidance, but this information can be recovered through an appropriate bridging strategy. A review of each combination can determine an appropriate decision rule for assignment of each combination to a single racial or ethnic category. In each case the assignment

<sup>26</sup> Department of Education, “proposed Guidance...,” p. 59272.

<sup>27</sup> Department of Education, “Final Guidance...,” p. 44870.

<sup>28</sup> See Broh and Minicucci, 2003, pp. 15-22.

complies with the requirements of the Guidance as long as it is documented, justified, and explained, such as keeping this paper on file.

Latino-White. This category is the largest combination of two race and ethnicity categories and likely results from the many years of contradictory instructions to the Latino community. Prior to the 1970 census, for example, anyone who called themselves “Mexican American” was counted as “White.” Latino now can be counted as any of the five race categories. As an example of the confusion, the population of Brazilian Americans may be relatively small, but they are concentrated in several northeastern urban areas with demographic characteristics similar to Latinos. The Guidance defines this group as White;<sup>29</sup> on the other hand, Latino is defined as including people with ancestry from South America.

Additionally, the implementation of the Guidelines in many collection formats will encourage or even require respondents who have chosen “Latino” in the first question of a two-question format to choose one of the additional five racial categories. Survey experiments in the mid 1990s showed that many Latinos preferred to skip the second question altogether or to select “Other” when it was available as an alternative.<sup>30</sup> The CPS experiments considered this combination as “lost information,” while an alternative interpretation is a “racialized” identity among Latinos.<sup>31</sup>

This combination shows up as “Latino” in IPEDS counts since Latino trumps any race. A bridging strategy, too, should count this combination as “Latino,” which will result in no differences with the IPEDS reporting.

Asian-White. The second largest two-race category will be lost to government reporting since these respondents will be reported in the “two or more races” category. Anecdotally, Asian students are likely to be undercounted because of the belief that affirmative action policies discriminate against the Asian population<sup>32</sup> This encourages selection of a second race category or no selection at all. Furthermore, inter-marriage among whites and non-whites is also greater among Asians than any other race. One out of five Asian women in the United States is married to a white man. The percentage of U.S. born Asian women is twice as high, 41 percent, and 30 percent for U.S. born Asian men married to white women.<sup>33</sup> Further, application data suggest that the number of Asian-White students is likely to grow in the future. COFHE guidelines call for this

<sup>29</sup> See. Helen Marrow, “To Be or Not To Be (Hispanic or Latino),” *Ethnicities*, Vol. 3, No. 4 (2003), pp. 427-464.

<sup>30</sup> Census 2010 is required by federal legislation to include a collection category called “other.”

<sup>31</sup> See Tienda and Mitchell, *op cit*.

<sup>32</sup> There is some evidence to support this common view. Espenshade, Thomas j. and Chang Y. Chung, “Admission Preference for Minority Students, Athletes, and Legacies at Elite Universities,” *Social Science Quarterly*, Vol. 85, No. 5 (December, 2004), pp. 1422-1446. David R. Colburn, Charles E. Young, and Victor M. Yellen. (2008). Admissions and Public Higher Education in California, Texas, and Florida: The Post-Affirmative Action Era. *InterActions: UCLA Journal of Education and Information Studies*. Vol. 4, Issue 1, Article 2.

<sup>33</sup> U.S. Census Bureau, “America’s Families and Living Arrangements: 2006,” Housing and Household Economic Statistics Division, Fertility & Family Statistics Branch, Table A

group to be classified as “Asian.” This preserves the information about non-white, Asian, identity of the student.

Black-White. The historical interest in this category dates back at least to miscegenation laws of the Nineteenth century and Jim Crow laws of the early Twentieth century. Concerned about inter-racial marriage, the U.S. Census gathered information about “octoroons” (one-eighth black ancestry), “quadroons” (one-fourth black ancestry), and mulatto (one-half black ancestry) from 1890 through 1920. Beginning in 1930, the U.S. Census stopped using these categories for “mixed-race” and adopted what many describe as the “one drop of blood” rule where a combination of African American ancestry with white ancestry was classified as “Negro.”

However repugnant some may find the assumption that the smallest trace of one’s ancestry defines a person into any racial category, a body of scholarly literature affirms use of the “one drop of blood” in American society. Although that literature has its critics, college recruitment and affirmative action programs have largely used it, resulting perhaps administratively in increasing numbers of students from African American and white identities enrolling at selective colleges and universities over the past four decades.<sup>34</sup> Preserving information about this group of underrepresented minorities to understand affirmative action policies is a high priority for a bridging strategy.

American Indian-White. Classification as “American Indian” is conceptually different from all other categories of race and ethnicity. The 1997 Standards define American Indian as follows:

A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment [emphasis added].

The first part of the definition is consistent with other government categories that record the identity of the individual. Simply selecting a category is sufficient information to classify respondents as Hispanic or any race category -- except American Indian. “Tribal affiliation or community attachment” is also necessary to be classified as “American Indian.”

The Common Application addresses this need for additional information by asking college applicants that identify as “American Indian” on the form to supply the name of their “tribal affiliation” and to indicate their registration number with the tribe. Even with these additional requirements, admission deans report that this classification is the most likely category to be misinterpreted. The older nomenclature of “Native American” exacerbated the problem as some understood the word “Native” to mean that they were born in the United States.

Parents of students at COFHE schools who reported having children that were American Indian and White in a 2006 survey have other family characteristics more common

---

<sup>34</sup> Rimer and Arenson, *op cit.*

among whites than American Indians. “For example, of those selecting American Indian *only*, 29 percent have incomes of \$50,000 or less, compared to 15 percent for those who select *both* American Indian and white (white alone is 11 percent). At the other end of the income scale, about 10 percent of those identifying only as American Indian have incomes over \$150,000, while 26 percent of the American Indians who also identify as white have incomes at that level (the white only number is higher still, 41 percent). Analysis of the income data suggests that the median income of the Indian-plus-white group is 44 percent higher than the Indian-only group.<sup>35</sup>”

The 2007 COFHE Enrolled Student survey had an open-ended question that is also responsive to the assignment of students with American Indian and White identities. Following the citizenship and race/ethnicity question, we asked students to reply to the following request:

**Colleges and universities have an educational interest in creating a diverse student body. For this purpose, please indicate your ancestry, nationality, ethnic origin, or tribal affiliation in the space below.**  
(For example: do you think of yourself as Italian, Jamaican, African Am., Cambodian, Ethiopian, Norwegian, Dominican, French Canadian, Haitian, Korean, German, Lebanese, Polish, Navajo, Mexican, Puerto Rican, Taiwanese, Ukrainian, and so on.).<sup>36</sup>

Respondents were given free-form space to write their answer, which was coded using SPSS Text Analysis for Surveys. Each response was coded according to its reference to a racial or ethnic group. For example, the mention of a tribal affiliation or a more general phrase like “Native American” was coded as a reference to “American Indian.” Mention of a European nation or ancestry like “German” or “Anglo Saxon” or “Jewish” was coded as a reference to “white.” These free-field responses reflect an explication of the respondent’s racial and ethnic identity from the original data collection that is used to classify each student.

Students who identified as both American Indian and White in the Guidance collection format were more likely to mention a reference to being White than a reference to being American Indian. Only 26 percent of the codable mentions in the free-form question were an American Indian references while 62 percent of the codable mentions were White references. The remaining mentions were references to some other group, such as Latino, Asian, or other. Relevant to this discussion, only one-in four mentions of race among students with a combination of both American Indian and White identities explained their background using terms or phrases that referred to their American Indian ancestry.

We also analyzed the data according to the order in which the students gave either an American Indian or a White reference to their background. The assumption is that the

<sup>35</sup> Broh and Minicucci, p. 18.

<sup>36</sup> For methodological justification of this approach, see Kenneth Prewitt, “Racial Classification in America: Where Do We Go from Here?” *Daedalus* (Winter, 2005), pp. 1-13. Prewitt was the Director of the U.S. Bureau of the Census for the 2000 census.

first reference of a race or ethnicity is more salient than the second reference, which in turn is more salient than later references. For every order of the mentions of a race or ethnicity – from the first reference through the tenth reference – these respondents were more likely to describe their background as white than American Indian.<sup>37</sup>

To retain information about “American Indian” that is as close as possible to the Guidance definition,<sup>38</sup> this group will be coded as White.

Latino-Black. This relatively smaller category has great conceptual importance in a bridging strategy. Because the Guidance makes clear that a respondent answering “yes” to the first question in a two-question format is always counted as “Latino,” Latino would seem to supersede, or trump, Black. Additionally, the standards prior to 1997 used nomenclature for Whites and Blacks as “White (non-Hispanic)” and “Black (non-Hispanic).” The use of the parenthetical expression “non-Hispanic” also implied that Latino supersedes Black.

However, the Guidance has an implied accounting that is greater than 100 percent. Tallies that appear in standard statistical reporting only rarely break the races from the second question into all logical combinations of Latino and non-Latino. In the Institutional Research lexicon, Latinos are often reported as duplicated counts. Whether or not the Department of Education follows this convention in its peer analysis tool or web reports is important but not dispositive; the media and other secondary analysis of data rarely describe the classification system accurately, especially when comparing underrepresented minorities.

In 2004, the COFHE admission officers were asked how they classified individuals that had responded to both “Latino” and “Black” on applications that allowed more than one response. The query related to data exchanges about the race and ethnicity of applicants, admits, and matriculants at COFHE schools. The response of the Deans and Directors of admission was nearly unanimous that these students are counted as Black; this exercise continues to guide our decision about a bridging strategy that is sensitive to trend analysis.

Latino-Asian This relatively small group will be reported as Latino in the IPEDS classification, which retains information about underrepresented minorities and conforms to the trumping rule in the Guidance.

Asian-Hawaiian The roughly 400,000 Native Hawaiians have political and land use claims in Hawaii that predate the 1893 overthrow of the monarchy and Hawaii’s 1959 admission to statehood. The inclusion of the group as a racial and ethnic group is closely

<sup>37</sup> Data are not shown here but are available upon request from the authors.

<sup>38</sup> In a chapter about enumeration of American Indians in the 1990 census that uses the same methodology as the Guidance, William Peterson concludes, “the statistics are so ill based that they are virtually worthless. See Person, p., 112.

tied to the efforts of the state's Congressional delegation and state leadership for federal recognition of benefits and programs for Native Hawaiians.<sup>39</sup>

Presumably, one reason for grouping "Native Hawaiian" with "other Pacific Islander" is the identification of the indigenous populations from the United States and territories that are not in North America. However, Census Bureau reports, government studies, or media presentations rarely combine Native Hawaiians and other Pacific Islanders with American Indians in a single category of indigenous United States populations.

Prior to Census 2000, the data collection categories included "Asian and Other Pacific Islander," which also included Native Hawaiians. This category was split into two categories in the 1997 Standards to become "Asian" and "Native Hawaiian and other Pacific Islander." A ven diagram of the pre-1997 standards and the post-1997 standards would show that Native Hawaiians, as well as indigenous peoples from Guam, Samoa, or the Philippines, are in the overlapping category of two classifications. Prior to 1997, they were "grouped with "Asians;" after 1997 they are not longer grouped with Asians. Combining Asian with Native Hawaiian and other Pacific Islander groups preserves the meaning and trend line of the pre-1997 "Asian and Other Pacific Islander" category.

Asian-Black This relatively small group often receives national attention from the publicity surrounding professional golfer Tiger Woods, who once called himself "Cablinasian" for Caucasian, black, Indian, and Asian. Interestingly however, he has never allowed his name to be associated with the Multiracial movement. To retain information about underrepresented minorities, this group will be reported as Black.

American Indian-Black This classification is historically interesting due to a myth that prevails in the African American population about a common ancestry with American Indians. A population of Black Seminole Indians as well as an infamous history of slave ownership among American Indians in the Carolinas provides evidence for the myth. DNA tests of African Americans, however, suggest that very few African Americans, in fact, have ancestors who were American Indian.<sup>40</sup> This group will be classified as African American.

Latino-American Indian This group will be classified as Latino for many of the same reasons described with other combinations of American Indians, as well as the Guidance procedure that answering the first question affirmatively classifies a person as Latino, regardless of responses to the second question.

Latino-Hawaiian This small group will be classified as Latino since those answering Hawaiian are very small in number and would have been combined with Asians in this reporting format.

---

<sup>39</sup> Janis L. Magin, "Occupation of Place Area Invigorates Native Hawaiian Movement," *New York Times*, May 3, 2008, p. A14.

<sup>40</sup> Gates, Henry Louis, "African American Lives 2," PBS series Black History Month, February, 2008.

American Indian-Asian This very small group will be classified as Asian for reasons that are similar to the combination of American Indian and White. On surveys, one might expect that this very small group has a large portion of erroneous entries.

Black-Hawaiian This population is conceptually very small and empirically one of the two smallest groups of students with two races at COFHE schools. Placing them into a “two or more” category or assigning them to either group will have little impact on a bridging strategy that differs from Federal reporting categories. Earlier discussion described the inclusion of Native Hawaiian’s and Asians as Asians. To preserve the black identity of those in this group, students should be included with others who are often the target of recruitment and affirmative action programs.

American Indian-Hawaiian This population too is conceptually very small and empirically the smallest group of students with two races in COFHE schools. Placing them into a “two or more” category or assigning them to either group will have little impact on a bridging strategy that differs from Federal reporting categories. Earlier discussion described the inclusion of Native Hawaiian’s and Asians as Asians. In COFHE’s study of open-ended questions where a respondent could describe an ancestry, only one person wrote “Hawaiian” and identified as American Indian but 32 identified as Asian. The expectation is actually that this combination is probably Native Hawaiian’s who confuse the word “American” to mean US citizenship and think of the word “Indian” to mean “original people.”

More than two This reporting category differs from the IPEDS reporting classification of “two or more” by excluding those who identify as Latino and one other race or who identify with only two races and are not Latino. More than two is, thus, a residual classification of those who identify with three, four, or five races as well as those who identify as Latino and two, three, four, or five races. Linguistically, it also describes the concept of “multi-racial” more accurately than “two or more,” which includes “bi-racial.”

The fifteen combinations, plus the multi-racial category of “More than Two,” can be summarized with the following set of decision rules for coding race and ethnicity.<sup>41</sup>

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 citizen question as “Foreign.”
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.

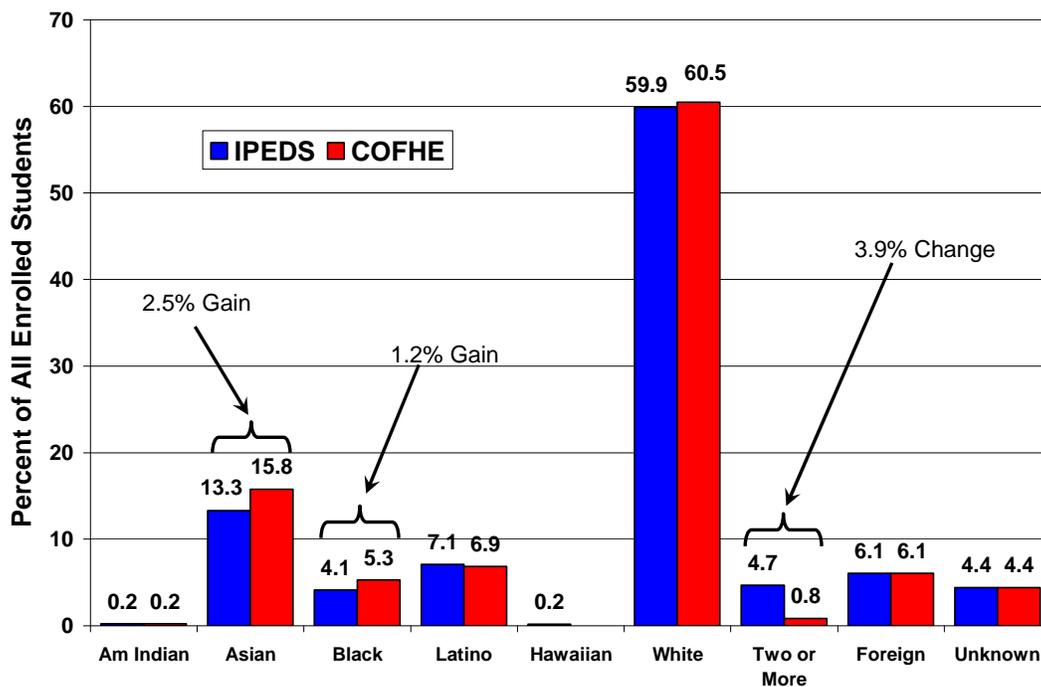
---

<sup>41</sup> See the SPSS code in Appendix A.

7. Hispanic supersedes identities other than Black.
8. Asian supersedes White
9. American Indian identities do not supersede any other category.

Figure 3 displays the results of these principles for assigning each category of students with two racial or ethnic identities into a single race or Hispanic category. The difference between the trumping rules for IPEDS reporting and the trumping rules for COFHE reporting are illustrated by the height of the bars in each reporting category. Note that the “Two or More” category for IPEDS classification is compared with “More than Two” for COFHE classification. Often conceptualized as “Multiracial,”<sup>42</sup> this category has the biggest

**Figure 3: Comparison between IPEDS and COFHE Reporting Format**



change among all of the IPEDS reporting categories: a decline of 3.9 percent. The decrease, however, is a reduction in a category with the least amount of information about race and ethnicity since it includes a mixture of different combinations of student identities. In this sense, it is a residual category for both the IPEDS and the COFHE designation, except to the extent that students identify as “multiracial” as a collective

<sup>42</sup> For a summary of this movement, see Kim W. Williams, *Mark One or More*. Ann Arbor: The University of Michigan Press, 2006.

concept rather than a combination of the race or ethnicity groups that they actually selected.<sup>43</sup>

Information gains are largely seen in the Asian and the Black categories. The COFHE reporting format has a 2.5 percent gain in information about Asian students and a 1.2 percent increase in information about Black students. While these numbers appear to be relatively small adjustments, they reflect an important percentage of students of color with these identities. The COFHE format recovers 18.8 percent of the information about Asian students and 29.3 percent of the information about Black students.

This gain in information is accomplished with an insignificant reduction in the percentage of Latino students, coming entirely from the assignment of students with a combination of Latino and Black identities to a Black classification. The trumping rules for all other combinations of Latino for COFHE reporting is consistent with the trumping rules for IPEDS reporting.

Additionally, the COFHE reporting format does not simply reassign majority students to some other category. Indeed, the percentage of white students actually increases slightly with COFHE reporting compared to IPEDS reporting. This is accomplished primarily through the recognition that the combination of American Indian and White includes students that largely have characteristics suggesting little, weak, or no tribal affiliation.

In sum, 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 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. The COFHE reporting format also retains important information about students of color. The Guidance recommends a bridging strategy for situations where the IPEDS format would not adequately reflect the administrative needs or the historical trends of the school. The COFHE format is an alternative that complies with the strict standards for collection in the Guidance and for the standards of reporting in the Guidance when it is accompanied with an explanation and documentation that is available for review. The COFHE reporting format could not be used for reporting for IPEDS but its information gains are useful for reporting to the public and for recording administrative objectives with affirmative action programs.

## SUMMARY AND CONCLUSION

The "Final Guidance, on Maintaining, Collecting and Reporting Racial and Ethnic Data to the Department of Education" imposes a collection format on higher education that

---

<sup>43</sup> However, Kim DaCosta interviewed multiracial activists who object to the current classification in the 1990s and finds that multiracialism is itself becoming an identity. On the other hand, her subjects also describe an existential conflict with virtually any classification, such as class, education, family, or sexuality. See Kimberly DaCosta, *Making Multiracials*. Palo Alto, CA: Stanford University Press, 2007.

many believe is flawed. However, colleges and universities must follow the requirements since Title IV monies and other federal funds are tied to compliance with the regulations. Adapting the format and nomenclature for data collection is chronologically the first step toward compliance. Since many institutions of higher education will need to modify paper forms and electronic systems for data collection, implementation of the Guidance is an opportunity to gather additional information that colleges and universities find useful in administrative processes. With this in mind, the Consortium on Financing Higher Education, the Common Application, and the College Board collaborated on a collection format that meets many admission and data collection needs while preserving the storage and reporting requirements for the Department of Education and other administrative objectives.

The data storage requirement for many schools will mean the creation of new data tables in legacy systems and adoption of additional coding schemes to store new information. Many commercial software vendors are distributing updates and revised programs to meet these needs. The addition of nationality and ethnic codes can be accomplished with the use of codes that do not necessarily require additional data fields in legacy systems. Schools are not required to resurvey their students although those that do so may map old racial and ethnic codes as default values into new values that comply with the storage and reporting requirements of the Guidance.

The reporting practices in the Guidance are also flawed and will result in the loss of information about race and ethnicity, especially Asians and African Americans, that is documented in this paper. An institutional reporting practice that preserves the information is clearly desirable to the reporting standards in the Guidance. Furthermore, the Guidance itself requires colleges and universities to store data for alternative reporting formats and it encourages colleges and universities to develop formats that meet their administrative and historical needs.

This paper provides documentation and justification for an alternative reporting strategy that retains and displays information that will go unreported in the IPEDS format. This effort is consistent with efforts of other organizations that engage in the collection of racial and ethnic data that are important to admission procedures. Because data about students typically enters administrative systems through admissions, the initiation of the information from the prescribed collection format can inform the data requirements for other offices and institutional storage and reporting.

The reporting format in this paper meets the requirements of the Guidance since institutions can use bridging strategies to produce reports according to multiple formats. Indeed the nomenclature in the COFHE reporting format are in one-to-one correspondence with the categories of the IPEDS reporting format.<sup>44</sup> The difference in frequencies for each category results from alternative trumping rules.

---

<sup>44</sup> The only categories with different nomenclature are “Two or more” in IPEDS and “More than Two” in COFHE. On the other hand, “More than Two” is a subset of “Two or More” and could be reported with this nomenclature to conform to the requirements of the Guidance even though it would logically have fewer cases.

Consequently, the COFHE rules should not and can not be used for submission of IPEDS data. However, with appropriate documentation on file, such as this paper, an institution could produce a report using the COFHE trumping rules for both internal and external usage. Over time, one would predict that either the Department of Education would alter its reporting format to meet institutional needs, that institutions would find the need for the additional information less useful than many institutional researchers currently believe, or that IPEDS reporting format would become a less trusted count of diversity than some other reporting format.

## BIBLIOGRAPHY

Broh, C. Anthony, "Summary: Provisional Guidelines on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity," *AIR Currents* 38 (Summer, 2001), p. 6.

Broh, C. Anthony, "Race to the Unknown," *Transcript*, An online News Source for members of the American Association for Collegiate Registrars and Admission Officers, March 22, 2007

Broh, C. Anthony, Correspondence with Patrick J. Sherrill, Department of Education in behalf of the Consortium on Financing Higher Education, September 22, 2006.

Broh, C. Anthony and Stephen D Minicucci, "One Organizations Bridging Strategy for Racial and Ethnic Classification," Paper presented at the Annual Forum of the Association for Institutional Research, Tampa, FL, May 21, 2003.

Colburn, David R., Charles E. Young, and Victor M. Yellen. (2008). Admissions and Public Higher Education in California, Texas, and Florida: The Post-Affirmative Action Era. *InterActions: UCLA Journal of Education and Information Studies*. Vol. 4, Issue 1, Article 2. <http://repositories.cdlib.org/gseis/interactions/vol4/iss1/art2>.

Department of Education, "Proposed Guidance on Maintaining, Collecting and Reporting Data on Race and Ethnicity to the U.S. Department of Education; Notice," *Federal Register*, Vol. 71, No. 151 (August 7, 2006), pp. 44866-44871.

Department of Education, "Supporting Statement for IPEDS 2007-2010, OMB Paperwork Reduction Act Submission," May 4, 2007.

Department of Education, "Final Guidance, on Maintain, Collecting and Reporting Racial and Ethnic Data to the Department of Education," *Federal Register*, Vol.72, No. 202 (October 19, 2007), pp. 59266-59279.

Equal Employment Opportunity Agency, "Agency Information Collection Activities: Notice of Submission for OMB Review, Final Comment Request," *Federal Register*, Vol 70, No. 227 (November 28, 2005), pp. 71294-71303.

Gratz v. Bollinger, 539 U.S. 244 (2003); Grutter v. Bollinger, 539 U.S. 306 (2003).

Espenshade, Thomas j. and Chang Y. Chung, "Admission Preference for Minority Students, Athletes, and Legacies at Elite Universities," *Social Science Quarterly*, Vol. 85, No. 5 (December, 2004), pp. 1422-1446.

Marrow, Helen, "To Be or Not To Be (Hispanic or Latino)," *Ethnicities*, Vol. 3, No. 4 (2003), pp. 427-464.

Katz, William Loren, "Black Indians," Africanamericans.com.  
<http://www.africanamericans.com/BlackIndians.htm>.

Magin, Janis L., "Occupation of Place Area Invigorates Native Hawaiian Movement," *New York Times*, May 3, 2008, p. A14.

Marrow, Helen, "To Be or Not To Be (Hispanic or Latino)," *Ethnicities*, Vol. 3, No. 4 (2003), pp. 427-464.

Office of Management and Budget, "Directive No. 15: Race and Ethnic Standards for Federal Statistics and Administrative Reporting," May 12, 1977.

Office of Management and Budget, "Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity; Notice," *Federal Register*, Vol. 66, No. 10 (January 16, 2001), pp. 3830-3831.

Peterson, William, *Ethnicity Counts*. New Brunswick: Transaction Publishers, 1997.

Prewitt, Kenneth, "Racial Classification in America: Where Do We Go from Here?" *Daedalus* (Winter, 2005), pp. 1-13.

Riles, Karen, "The Seminole-Indian Negro Scouts" Texas Indians,  
<http://www.texasindians.com/bsem.htm>.

Rimer, Sara and Karen W. Arenson, "Top Colleges Take More Blacks, But Which Ones?" *The New York Times*, Education Section, April 21, 2008.

Tienda, Marta and Faith Mitchell [ed.], *Hispanics and the Future of America*, Report of the Panel on Hispanics in the United States, Committee on Population, Division of Behavioral and Social Sciences and Education. Washington, DC: National Research Council, 2006

U.S. Census Bureau, "America's Families and Living Arrangements," 2006 Current Population Survey, Housing and Household Economic Statistics Division, Fertility & Family Statistics Branch, 2006, Table A.  
<http://www.census.gov/population/www/socdemo/hh-fam/cps2006.html>

Williams, Kim W., *Mark One or More: Civil Rights in Multiracial America*. Ann Arbor: The University of Michigan Press, 2006.

## APPENDIX A

### SPSS CODE FOR THE COFHE REPORTING FORMAT

```
RECODE asian black natam hisp white hawaii (SYSMIS=0).
EXECUTE.
```

```
COMPUTE hawasian = 0.
IF (asian = 1 or hawaii = 1) hawasian = 1.
COMPUTE racecount = hawasian + black + natam + hisp + white.
```

```
DO IF (citizen > 2).
COMPUTE cofherace = 7.
ELSE IF (racecount > 2).
COMPUTE cofherace = 6.
ELSE IF (black = 1).
COMPUTE cofherace = 3.
ELSE IF (hisp = 1).
COMPUTE cofherace = 4.
ELSE IF (asian = 1 or hawaii = 1).
COMPUTE cofherace = 2.
ELSE IF (white = 1).
COMPUTE cofherace = 5.
ELSE IF (natam = 1).
COMPUTE cofherace = 1.
ELSE.
COMPUTE cofherace = 8.
END IF.
```

```
EXECUTE.
```

```
VARIABLE LABELS cofherace 'Race according to Tony guidelines'.
VALUE LABELS cofherace
  1 'American Indian or Alaska Native'
  2 'Asian, Native Hawaiian, or Other Pacific Islander'
  3 'Black or African American (Not Hispanic or Latino)'
  4 'Hispanic or Latino'
  5 'White (Not Hispanic or Latino)'
  6 'More than Two'
  7 'Foreign'
  8 'Unknown'.
```

```
FREQUENCIES  
VARIABLES=cofherace  
/ORDER= ANALYSIS .
```