

Suite 402
238 Main Street
Cambridge, MA 02142-1046
(617) 253-5030
FAX (617) 258-8280

Consortium On Financing Higher Education

September 22, 2006

Patrick J. Sherrill
U.S. Department of Education
400 Maryland Avenue, SW
Room 6C103
Washington, DC 20202-0600

Member Institutions

Amherst College
Barnard College
Brown University
Bryn Mawr College
Carleton College
Columbia University
Cornell University
Dartmouth College
Duke University
Georgetown University
Harvard University
The Johns Hopkins University
Massachusetts Institute
of Technology
Mount Holyoke College
Northwestern University
Oberlin College
Pomona College
Princeton University
Rice University
Smith College
Stanford University
Swarthmore College
Trinity College
The University of Chicago
University of Pennsylvania
The University of Rochester
Washington University
in St. Louis
Wellesley College
Wesleyan University
Williams College
Yale University

Dear Mr. Sherrill:

The Secretary is requesting comment on “Guidance for Data on Race and Ethnicity.” Having won the right to consider race as a factor in admissions in the Supreme Court case of *Gratz v. Bollinger*, colleges and universities have a large stake in knowing that the methods for collecting, maintaining, and reporting racial and ethnic data reflect their efforts at fairness in using race as a factor in admissions. While IPEDS data do not ask for a racial/ethnic breakdown of admission statistics, enrollment data at colleges and universities are typically collected from applications and materials related to the admission process. As such, the guidelines should allow simple, authentic, and truthful self-identification as well as reproducible measurement of information in the application and admission to college.

I engage in analytic work on student enrollment data for a consortium of U.S. colleges and universities. These schools are concerned that changes in data collection protocols will have an adverse consequence in their enrollment management efforts. The objective of my letter is to identify shortcomings in the proposed standards and to recommend changes that improve the quality of the data. I provide argument and evidence from research data to support three recommendations:

Recommendation 1: Use the “one-question format” rather than the “two-question format” for collecting data on race and ethnicity.

Recommendation 2: Allow a “bridging strategy” for comparing current data to pre-2000 data to govern the IPEDS report format.

Recommendation 3: Add Question wording that explains the context for gathering data on race and ethnicity.

The “two-question format” that is currently required in the standards is a departure from prior government and institutional practices and does not serve the needs of higher education nor the data collection requirements in the Department of Education.

However, the opportunity for institutions of higher education to use an institutional methodology for a so-called “bridging strategy” to compare current and historical data should also be acceptable for IPEDS reporting.

RECOMMENDATION 1: USE THE “ONE-QUESTION FORMAT,” RATHER THAN THE “TWO-QUESTION FORMAT,” FOR COLLECTING DATA ON RACE AND ETHNICITY.

The requirement to collect data according to the proposed “two-question format” creates a restriction that does not meet institutional objectives nor is it necessary for the Department to carry out its responsibilities. As explained in the proposed Guidance,

Using the two-question format, the first question asks whether or not the respondent is Hispanic/Latino. The second question allows individuals to select one or more races from five groupings: (1) American Indian or Alaska Native (2) Asian (3) Black or African American (4) Native Hawaiian or other Pacific Islander (5) white.

The proposed Guidance explicitly forbids use of a “one-question format” that would have the following six racial/ethnic identities listed together (1) American Indian or Alaska Native (2) Asian (3) Black or African American (4) Hispanic/Spanish origin or Latino (5) Native Hawaiian or other Pacific Islander (6) white.

The justification for the requirement is the claim that the “two-question format “results in more complete reporting of Hispanic ethnicity.” The concern I raise is based on the following observations:

1. *The requirement to use the two-question format is not based on scientific research.* I know of no research that has been conducted by any government agency that substantiates the claim that the “two-question format” produces “more complete reporting.” When a citation for this conclusion was requested, the research from the 1995 CPS Supplement was cited as authoritative. However, this analysis was not conducted with instructions that allowed respondents to select “one or more” racial identities. No *bona fide* experimental research has been provided to substantiate the claim that the two-question format is more accurate.
2. *Secondary analysis of the 1995 data suggests that the two-question format did not produce more accurate data than the one-question format.* The 1995 data are instructive in observing what respondents do with a two-question format and a one-question format even though the purpose of the experiment only addressed the order of the Hispanic and race questions and whether a category of “other” made a

difference in responses. Our review of the data¹ suggests that there is very little, if any, difference between the two-question format and the one-question format in the percentage of respondents who identified themselves as “Hispanic” or other underrepresented minorities.

3. *The one-question format provides the same opportunity for respondents to identify as Hispanic and/or with any race.* The fact that both formats allow for exactly 63 race/ethnic combinations means that no information is lost in the use of the “one-question format.” Additionally, the “one-question format” and the “two-question format” have identical storage requirements in information systems.
4. *The 2006 panel at the National Academy of Sciences on Hispanics and the Future of America² provides evidence that the two-question format was confusing to the young Hispanic population.* The primary findings of the NAS study was that the “Hispanic-Latino” category that was initially introduced into the census as an administrative category was transformed into an ethnic category and has become a “biological marker of social difference” (p. 40) equivalent to a race category. Indeed, the authors of the Report produce survey data of Hispanic families to show that the “children of immigrants learn to see themselves more and more in these terms – as members of a racial minority [emphasis added] —and even to racialize their national origin” (p. 41). In another section of the Report the authors present the results of an open ended question where Hispanic respondents were asked to identify their “race.” Only 6.6 percent wrote “black” and 3.8 percent chose “white;” the others wrote “Hispanic” or “Latino” or a variety of mixed “Latino” and national origin answers. The National Academy of Sciences concludes: “Indeed, ‘Hispanic’ was the label most consistently given by the respondents to characterize their own *racial* [emphasis in original] identity....” (p 43)
5. *The two-question format distorts the true identity of “underrepresented minorities.”* Philosophically, this format says, “we care more if you indicate that you are Hispanic than if you indicate you are black or American Indian, etc.” Separating the identities of Hispanics from other groups is a visual statement that groups are not treated equally in higher education policy. (While the reporting format in the past has implied a hierarchy in the use of nomenclature such as “white (non-Hispanic)” or “black (non-Hispanic),” the collection format has given respondents an equal opportunity to indicate their identity. Equality of opportunity should extend to the data collection format that is used to measure educational progress.

What is required for Hispanics is not allowed for other underrepresented minorities. For example if a respondent identifies as a combination of Hispanic and black, the person’s identification is reported as Hispanic; if the respondent identifies as a combination of white and black, the person identification is reported as “Two or

¹ Research note available from the COFHE upon request.

² 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.

More.” Thus only some combinations of underrepresented minorities (e.g., Hispanic and black) get reported while other combinations (e.g., blacks and American Indians, blacks and Native Hawaiians) don’t get reported.

But even if this reporting requirement stands the test of time or is modified in institutional or future government statistics, the two-question format distorts the underrepresented minority count by emphasizing the importance of Hispanic groups, more than other underrepresented minorities, to indicate their identity in a separate question.

6. *Requiring higher education to use the two-question format is a major departure from prior practices.* OMB Statistical Directive 15, adopted in 1977, recognized the need for flexibility by allowing government agencies and reporting institutions to select either the one-question format or the two-question format. The Secretary recognized that different policy objectives required optional formats. Nothing has changed since that time except the opportunity for respondents to select more than one race, and this change invites, rather than restricts, the use of the one-question format as a statement of equal opportunity. The Office of Management and Budget reaffirmed this conviction in 1997 and 2001 by only “recommending,” but not requiring, the two-question format for sectors of the economy where the one-question format was more appropriate; OMB specifically mentioned higher education as a sector where an alternative might be appropriate. Indeed, even the 2005 EEOC guidelines still allow data collection using the one-question format.

Historically institutions of higher education have chosen to exercise their option under OMB Statistical Directive 15 with use of the one-question format. College admission applications and student registration forms throughout the country almost universally have selected the one-question format for the past three decades. Indeed many colleges and universities have adopted most of the changes in the proposed standards, including the choice of one or more races, for several years, but they have chosen the one-question format because of the implied equality that the institutions wished to communicate to students, the greater simplicity in structuring the question on forms, and the greater understanding among students about their options for identifying with a racial or ethnic group.

7. *The two-question format is unnecessary for maintenance and reporting requirements.* All of the combinations of race and ethnicity that are required in the current standards, as well as the implied computer requirements for storing dichotomous indicators of race and ethnicity, can be accommodated by the one-question format.

RECOMMENDATION 2: ALLOW THE “BRIDGING STRATEGY” FOR COMPARING CURRENT AND PRE-2000 DATA TO GOVERN THE IPEDS REPORTING FORMAT.

The proposed standards allow educational institutions to use a “bridging strategy” for reporting data to make calculations comparable to pre-2000 data collection when “mark one or more” was typically not an instruction to respondents. We applaud this flexibility in the standards. We believe that many institutions will utilize the option to report a “bridging strategy” so they can address both the educational objectives and the public policy issues that affect their institution. In particular, the desire to have racial diversity on campus motivates many admission offices to collect racial and ethnic data for comparison over time in order to monitor progress of affirmative action policies.

For several years, most institutions have used the “mark one or more” instruction for a single list of race and ethnic categories. These schools have grappled with the problem of how to report data for comparison to previous years – many of which go back three decades to OMB Statistical Directive 15. The institutional determination of reporting categories for multiple responses is considerably different from the reporting decisions for reporting multiple responses in IPEDS. The following table highlights some of the differences between the Department requirements and common reporting practices in higher education:

COLLECTION response	IPEDS REPORTING requirement	COMMON REPORTING practices³
Neither “yes” nor “no” to “Are you Hispanic” in a two-question format <u>and</u> the choice of a single race	Other, unknown	The selected race
Black and Hispanic	Hispanic	Black
Black and White	Two or more	Black
Black and Asian	Two or more	Black
American Indian and Hispanic	Hispanic	American Indian
American Indian and white	Two or more	American Indian
Native Hawaiian and Asian	Two or more	Asian
Native Hawaiian and white	Two or more	Native Hawaiian

Institutions that make a good faith effort to assign combinations of race and ethnicity into a single category for comparison to pre-2000 data should be allowed to use these same decision rules in reporting IPEDS data. Consistency between the institutional calculations and the IPEDS reporting format creates verifiable information, less regulatory burden, and

³ 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.

data that are consistent with the public policy objectives of the institution. We believe the proposed Guidance allow this flexibility and should be clarified and supported in the final Guidance.

RECOMMENDATION 3: ADD QUESTION WORDING THAT EXPLAINS THE CONTEXT FOR GATHERING DATA ON RACE AND ETHNICITY

Providing the reason for the data collection will ensure that students understand why colleges are seeking information about race and ethnicity in a college application and related materials. Many applicants are unfamiliar with the history of affirmative action, the scholarly research and legal briefs about the benefits of racial and ethnic diversity in an educational setting, or the opinion of the Supreme Court in allowing the use of race as a factor in admissions. A brief one sentence explanation provides context for the data collection.

In the first section of this letter, we discussed the problems with the proposed “two-question format” for collecting data on race and ethnicity from students who are entering higher education. Using the “one-question format,” in which “Hispanic, Spanish origin, or Latino” is interspersed with the five races, we propose the following wording for the collection of racial and ethnic data:

Colleges and Universities are asked by many, including the Federal government, college guides, newspapers and our own communities, to describe our students using many backgrounds. For this purpose, please check one or more of the following groups you consider yourself to be in.

- American Indian or Alaska Native**
- Asian American**
- Black or African American**
- Native Hawaiian or Other Pacific Islander**
- Hispanic, Spanish origin, or Latino**
- White**

Thank you for the opportunity to respond to the “Guidance for Data on Race and Ethnicity.” I look forward to discussing these and other proposals at the Technical Review Panel on November 14-15.

Sincerely,



C. Anthony Broh
Director of Research Policy

Cc: Kristine E. Dillon, President