

**AERA Annual Meeting 2001
Seattle, WA**

**Legal Challenges to High-Stakes Testing:
A Case of Disparate Impact in Michigan?**

**Donald E. Heller
Douglas T. Shapiro**

**Center for the Study of Higher and
Postsecondary Education**

University of Michigan

Michigan Merit Award Scholarship Program

- Enacted June 1999, with goal of “increas[ing] access to postsecondary education and reward[ing] academic achievement”
- Funded from state’s tobacco settlement funds
- One-time \$2,500 scholarship awarded based on:
 - Passing all four 11th grade MEAP tests
 - Taking all four MEAP tests, passing two, and scoring in top 25% nationally on SAT, ACT, or WorkKeys test
- Federal lawsuit filed June 2000 alleging disparate impact and violation of 14th Amendment and CRA rights of minority students

Questions in This Study

- What is the relevant case law and theoretical frameworks relating to testing in educational settings?
- What is the racial distribution of the first cohort of students eligible for the Michigan scholarships, and is there *prima facie* evidence of disparate impact?
- Is the scholarship program likely to meet the legislated goal of increasing access to higher education in Michigan?

Legal Framework – Disparate Impact

- A disproportionate disadvantage imposed on a group of students from the application of a facially neutral criterion or test, where the criterion is not justified by educational necessity.
- Disparate impact claims have been brought under:
 - Equal Protection Clause of the 14th Amendments
 - Title VI of the Civil Rights Act of 1964
 - Title IX of the Education Amendments of 1972
 - Title II of the Americans with Disabilities Act of 1990
 - §504 of the Rehabilitation Act of 1973
- No requirement to show discriminatory intent

Proving Disparate Impact

- Three steps:
 1. Plaintiff must show that the practice has a disproportionate effect, using one of two tests:
 - Statistical significance: Success rates among groups are different from a random distribution
 - 4/5 or 80% test: Success of protected group is less than 4/5 of the most highly selected group

(Heubert & Hauser, 1999)

Proving Disparate Impact

2. Defendant may show a substantial legitimate justification, or “educational necessity,” including:
 - the test has been designed and validated as a measure of the intended construct
 - the specific cutoff scores employed (if any) bear a significant relationship to the educational competencies intended
3. Plaintiff may still prevail by showing that an equally effective alternative method exists and has a less discriminatory impact

Disparate Impact – Key Precedent

- *Sharif v. New York State Education Department* (S.D.N.Y. 1989)
- Federal district court barred the use of the SAT as a sole criterion for awarding merit scholarships
- Court found that female students had a reduced chance of winning the scholarship, without educational justification, in violation of Title IX

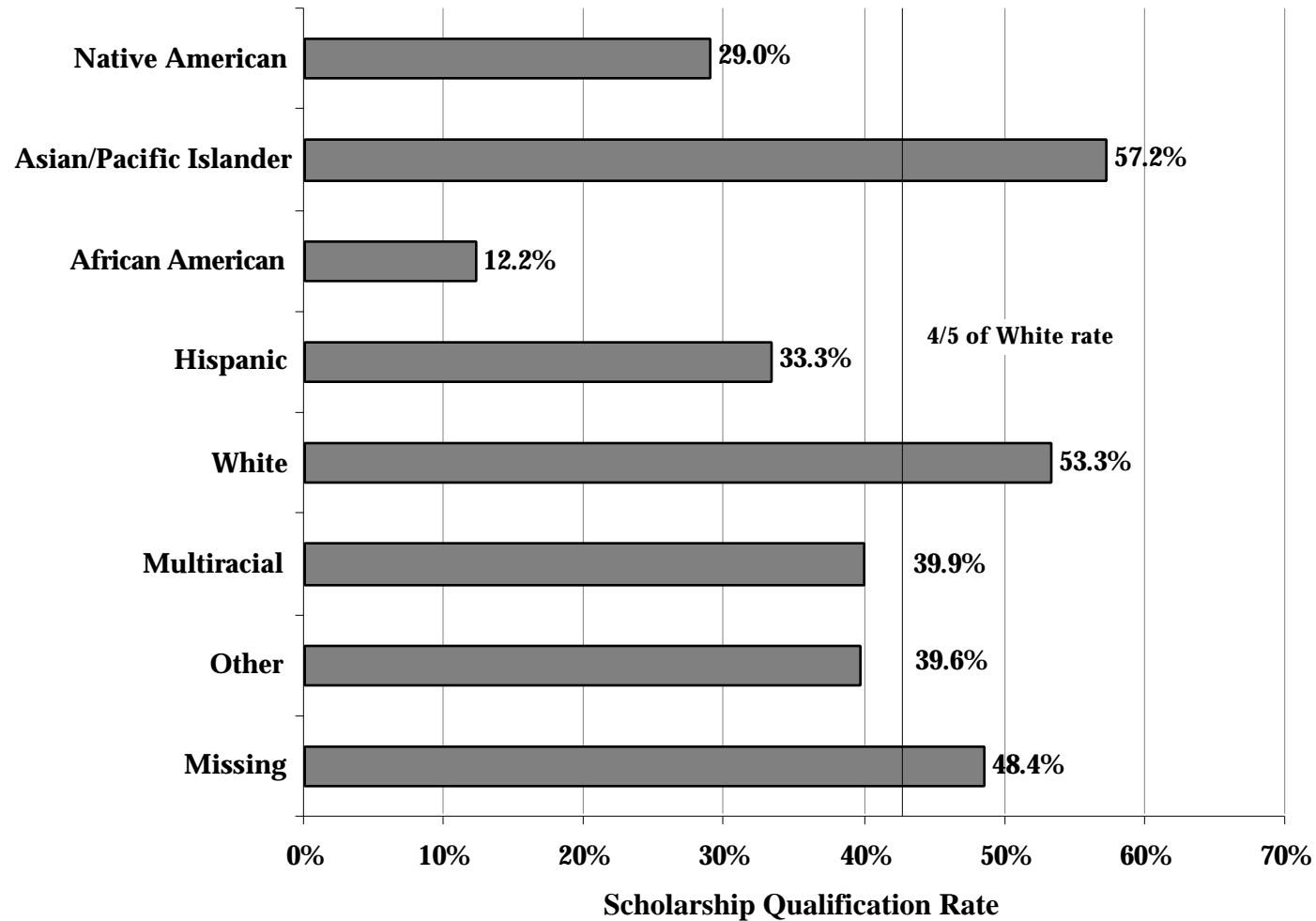
Testing and Socioeconomic Status

- Relationship between standardized test scores and student demographics is well-established:
 - Clear evidence that, nationwide, Whites and Asian Americans consistently score higher than African Americans, Hispanics, and Native Americans on all types of tests
 - High-SES students score higher than low-SES students

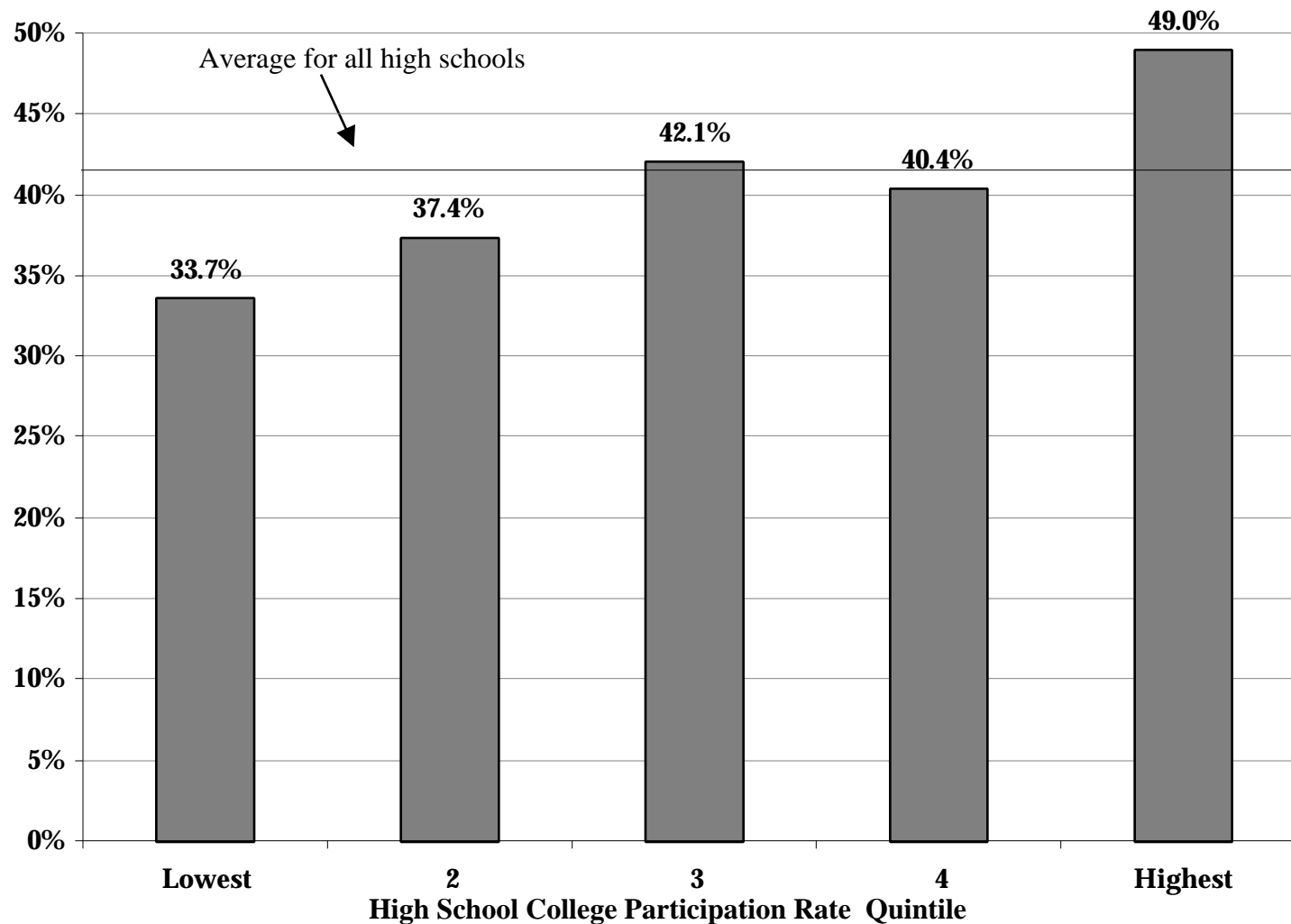
Research Methodology

- Bivariate analysis of the first cohort of high school graduates eligible for the Michigan Merit Scholarships
- Application of the statistical significance and 4/5 tests
- Comparison of characteristics of the first cohort of students with known patterns of college participation in Michigan

Results – Scholarship Qualification Rates by Race



Results – Rates by HS College Participation Rate



Discussion – Step 1. Disparate Impact

- Our results demonstrate that there are statistically significant differences in the rates at which students of different races qualify for the Michigan Merit Award Scholarships.
- African American, Hispanic and Native American groups are disproportionately underrepresented among scholarship qualifiers relative to their representation among test-takers
- These differences meet both the first and second legal tests for significance.

Discussion – Step 2. Educational Necessity

- Use of MEAP test for scholarship awards conflicts with AERA/APA/NCME standards on the use of testing in education
 - Not validated for purpose used (Standard 1.4)
 - MEAP test was never validated as a measure of individual student achievement
 - The test was designed not to measure individual performance, but to rate school district performance in meeting curricular standards
 - High-stakes decision is based on a single test (Standard 13.7)
 - 93.3% of scholarships were awarded via MEAP score alone
 - Receipt of scholarship can easily mean the difference between attending college or not

Discussion – Step 3. Alternative Methods

- Courts have recognized combinations of high school GPA and standardized test scores as less discriminatory alternatives to test scores alone for measuring academic achievement
- Most state merit scholarship programs use GPA or combination of GPA and test scores

Is the Program Increasing Access to Higher Education in Michigan?

- Our findings suggest that the scholarships are missing the mark:
 - The demographic groups with the lowest rates of participation in higher education are precisely those we identified as being the *least* likely to benefit from the scholarships – African American, Hispanic, Native American and low income students.
 - Conversely, the students to whom the scholarships are disproportionately awarded – Whites, Asian Americans and students in wealthier school districts – are least responsive to financial aid: they are most likely to attend college anyway, even without the grants.

For Further Details

- A copy of the paper will be available at:

<http://www.umich.edu/~dheller>

- A copy of the ACLU brief in *White et al. v. Engler et al.* may be downloaded from:

<http://www.aclumich.org/briefs/meap.htm>