7-31-2001

Does the LSAT Mirror or Magnify Racial and Ethnic Differences in Educational Attainment: A Study of Equally Achieving Elite College Students

William C. Kidder

Follow this and additional works at: http://scholarship.law.berkeley.edu/californialawreview

Does the LSAT Mirror or Magnify Racial and Ethnic Differences in Educational Attainment?: A Study of Equally Achieving "Elite" College Students

William C. Kidder†

TABLE OF CONTENTS

Does the LSAT Mirror or Magnify Racial and Ethnic Differences in Educational Attainment?: A Study of Equally Achieving "Elite" College Students

William C. Kidder

Affirmative action in higher education admissions is now prohibited in several states and more legal challenges are pending. This situation adds urgency to the question of whether "race-neutral" reliance on the LSAT is fair to students of color. In this comment, a database of law school applicants from fifteen highly selective institutions is used to test whether the LSAT produces racial and ethnic differences among students with equivalent performance in college after controlling for institution attended, grade-point average, graduation date and field of study. Based on the data and a review of the legal education literature, it is argued that the LSAT is culturally biased in the sense that it creates an artificial barrier to entering the legal profession. In other words, racial and ethnic gaps on the LSAT are found to be larger than differences in undergraduate grades, law school grades or measures of subsequent success in the legal profession. Causes of test bias, including stereotype threat in the testing milieu, are investigated. Next, recurring problems are examined regarding the traditional predictive validity approach to assessing fairness on the LSAT. Several alternative hypotheses for these findings are explored. It is concluded that test bias can be a rationale for justifying affirmative action programs, and that traditional law school admission criteria, which are driven by LSAT scores, should be significantly altered.

INTRODUCTION

A growing number of American law schools, and flagship public law schools in particular, have recently terminated race-conscious affirmative action plans in order to comply with various court decisions, popular referenda, and actions by public officials.[1] Moreover, it is quite possible that the United States Supreme Court will soon grant review to one of the several

---

1.   *See infra* Part 1.A.

pending challenges to affirmative action at public universities.[2] As affirmative action continues to come under fire, high-stakes standardized tests like the Law School Admission Test ("LSAT") have also become the focus of intensified criticism.[3] Much of the debate centers on whether standardized tests like the LSAT and the SAT are neutral barometers of racial and ethnic differences in educational achievement.[4]

In this Comment, I attempt to provide empirical answers to the question of whether students of color with the same undergraduate grades systematically score lower on the LSAT than white students, even when controlling for factors such as which college they attended and what undergraduate major they selected. I also compare differences in law school grades to differences in LSAT scores. This investigation of racial and ethnic differences in LSAT scores among college peers is relevant to policy makers, I argue, because it provides an important piece of evidence establishing cultural bias on the LSAT or in the testing milieu.

In this study, I matched African American, Chicano/Latino, Native American, and Asian Pacific American applicants with White applicants who possessed equivalent undergraduate grade-point averages ("UGPA") from the same colleges during the same time period. I relied on a database of 1996, 1997, and 1998 applicants from fifteen highly selective colleges and universities to Boalt Hall, the law school at the University of California ("UC"), Berkeley. This is the first attempt to replicate the UGPA-matching procedure developed in Joseph Gannon's 1981 pioneering study.[5]

The results indicate that among law school applicants with essentially the same performance in college, students of color encounter a substantial performance difference on the LSAT compared to their White classmates. These gaps are most severe for African American and Chicano/Latino applicants. A second round of matching, controlling for choice of major within each college or university, does nothing to reduce these performance differences on the LSAT. The results of this study therefore counter the claims of several standardized testing enthusiasts and affirmative action critics that the LSAT provides a neutral method of assessing academic achievement.[6]

Part I surveys recent controversies over affirmative action in higher education and introduces the present debate over whether standardized tests mirror or magnify differences in educational attainment. Part II

---

2. *See infra* Part I.A.
3. *See infra* Part I.A.
4. *See infra* Parts I.B.-C.
5. Joseph Gannon, *College Grades and LSAT Scores: An Opportunity to Examine the "Real Differences" in Minority-Nonminority Performance, in* TOWARDS A DIVERSIFIED LEGAL PROFESSION 272 (David M. White ed., 1981). Gannon's study is summarized *infra* Part IV.A.
6. *See infra* Part I.D.

describes the applicants included in the Boalt Hall database and details the matching methodology. Part III reports average differences in LSAT scores by race and ethnicity after comparing applicants with equivalent UGPAs in the same institution. It also describes the results obtained from a second round of matching where undergraduate major was added to these controls. Part IV examines Joseph Gannon's earlier research, explains why the present methodology was adopted over the conventional predictive validity approach, and explores three plausible alternative explanations for the substantial gaps on the LSAT among academic peers. I conclude that the LSAT systematically disadvantages minority law school applicants. I therefore argue that affirmative action can be justified as a corrective for those racial and ethnic biases that use of the LSAT introduces into the admissions process. I finally argue that it is essential to revamp admission criteria to reduce the influence of the LSAT, particularly at law schools that are prohibited from using race in admissions decisions.

I
BACKGROUND: STANDARDIZED TESTING AND THE POST-AFFIRMATIVE
ACTION LANDSCAPE

A.   The Current Status of Affirmative Action in Higher Education

Affirmative action in higher education is under greater threat today than at any time in the past two decades. In the wake of the Fifth Circuit's decision in *Hopwood v. Texas*,[7] California's Proposition 209,[8] the

---

7.   78 F.3d 932 (5th Cir. 1996) (ruling that the affirmative action program at the University of Texas ("UT") Law School was unconstitutional because taking account of race to improve the quality of learning was not a compelling governmental interest, and because the program was not narrowly tailored to remedy the UT Law School's own prior discrimination against minority students). The consequences of *Hopwood* at the UT Law School are discussed *infra* Part IV.C.2. *See also* William E. Forbath & Gerald Torres, *Merit and Diversity after Hopwood*, 10 STAN. L. & POL'Y REV. 185 (1999); Thomas D. Russell, *The Shape of the Michigan River as Viewed from the Land of* Sweatt v. Painter *and* Hopwood, 25 LAW & SOC. INQUIRY 507, 507-08 (2000). For a legal analysis of *Hopwood* from the perspective of a professor on the UT Law School defense team, see Samuel Issacharoff, *Can Affirmative Action Be Defended?*, 59 OHIO ST. L.J. 669, *passim* (1998). For a differing view, see Gabriel J. Chin et al., *Rethinking Racial Divides—Panel on Affirmative Action*, 4 MICH. J. RACE & L. 195, 201, 202 (1998) (criticizing UT's affirmative action plan as not defensible under a *Bakke* standard, and also criticizing their litigation strategy as one of "recreational litigation" that put "the lives of people of color on the line for no good reason").

8.   Proposition 209, now CAL. CONST. art. 1, § 31, was passed by a 54% majority of California's voters in November 1996. It states: "The State shall not discriminate against, or grant preferential treatment to, any individual or group on the basis of race, sex, color, ethnicity, or national origin in the operation of public employment, public education, or public contracting." For a description of the political fight over Proposition 209, see LYDIA CHAVEZ, THE COLOR BIND: CALIFORNIA'S BATTLE TO END AFFIRMATIVE ACTION (1998); NICHOLAS LEMANN, THE BIG TEST: THE SECRET HISTORY OF THE AMERICAN MERITOCRACY (1999). For a legal and policy analysis of Proposition 209, see Neil Gotanda, *Failure of the Color-Blind Vision: Race, Ethnicity, and the California Civil Rights Initiative*, 23 HASTINGS CONST. L.Q. 1135 (1996); David Benjamin Oppenheimer, *Carcieri's Self-Described "Progressive" Critique of the ACLU on Proposition 209: A "Conservative" Response*, 39 SANTA CLARA L. REV. 1153 (1999); Girardeau A. Spann, *Proposition 209*, 47 DUKE L.J. 187 (1997). For a

University of California Regents' SP-1 Resolution,[9] Washington's 1-200 Initiative,[10] Governor Jeb Bush's executive order replacing affirmative action with the "One Florida" plan,[11] and the chilling effect of threatened litigation,[12] a substantial number of America's leading public universities and

---

far-reaching and detailed analysis of racial mistreatment of people of color in California, and how these practices set the stage for the state's present educational inequities, see Richard Delgado & Jean Stefancic, *California's Racial History and Constitutional Rationales for Race-Conscious Decision Making in Higher Education*, 47 UCLA L. REV. 1521 (2000).

9.   The UC Regents approved SP-1 by a vote of 14 to 10 in July, 1995. SP-1 ended race-conscious admissions at the graduate and professional level beginning on January 1, 1997, a year before Proposition 209 took effect. *See* Cecilia V. Estolano et al., New Directions in Diversity: Charting Law School Admissions Policy in a Post-Affirmative Action Era app. B (May 9, 1997) (unpublished report, on file with Boalt Hall Law Library) (quoting SP-1 in its entirety). Proposition 209 applies to "preferential treatment" broadly, not just to admission decisions. Thus, Proposition 209 places limitations on financial aid and scholarships, whereas SP-1 does not. *See* Robert Cole et al., Report of an Ad Hoc Task Force on Diversity in Admissions, University of California, School of Law, Berkeley 114 (Oct. 14, 1997) (unpublished report, on file with Boalt Hall Law Library). For an analysis of the consequences of SP-1 and Proposition 209 on racial diversity at University of California law schools in the three years since the ban on affirmative action, see William C. Kidder, *Situating Asian Pacific Americans in the Law School Affirmative Action Debate: Empirical Facts About Thernstrom's Rhetorical Acts*, 7 ASIAN L.J. 29 (2000) (analyzing 1994-99 admission data for Boalt Hall, UCLA, and UC Davis law schools) [hereinafter Kidder, *Situating Asian Pacific Americans*]; Jerome Karabel, The Rise and Fall of Affirmative Action at the University of California tbl.11 (Sept. 1999) (unpublished manuscript, on file with the UC Berkeley Institute for the Study of Social Change) (calculating that for the 1997 entering classes, the ban on affirmative action turned back the clock thirty years at Boalt and twenty-nine years at UCLA Law School in terms of the last time there were comparably small proportions of Blacks and Latinos); Nanette Asimov, *Asian American Students Trickle in to Law School*, S. F. CHRONICLE, Feb. 23, 2001 at A2 (reporting that Proposition 209 has not boosted Asian American enrollments at UC law schools, contrary to earlier claims). Recently, the UC Regents voted to overturn SP-1. Tanya Schevitz, *Affirmative-Action Ban Revoked by UC Regents*, S.F. CHRON., May 17, 2001, at A1.

10.   The voters of Washington passed Initiative 200, a ballot initiative with wording identical to Proposition 209, in November 1998 with a 58% majority. *See, e.g.*, D. Frank Vinik et al., *Affirmative Action in College Admissions: Practical Advice to Public and Private Institutions for Dealing with the Changing Landscape*, 26 J.C. & U.L. 395, 413-15 (2000). For a description of admissions at the University of Washington post-Initiative 200, see Siobhan Gorman, *After Affirmative Action*, NAT'L J., Apr. 8, 2000, at 1120; *Test Score Use and Affirmative Action Bans*, FAIRTEST EXAMINER, Spring 2000, at 12 (reporting that at the University of Washington the number of African American freshman in 1999 was down 42% compared to 1998).

11.   The "One Florida" plan, adopted in November 1999 by executive order, discontinues race-conscious affirmative action in the University of Florida system. Although the race ban is also in effect at the graduate and professional level, there is no analogous admissions plan for law, medical, business, and graduate schools. For information on the development and projected consequences of the "One Florida" plan, see *Why the "One Florida" Plan Would Remove Blacks from the Best Campuses of the University of Florida*, 27 J. BLACKS HIGHER EDUC. 29, 29-30 (2000); Jeffrey Selingo, *What States Aren't Saying About the 'X-Percent Solution'*, CHRON. HIGHER EDUC., June 2, 2000, at 31. Governor Bush's executive order was, in part, an effort to fend off an anti-affirmative action ballot initiative championed by UC Regent Ward Connerly. *See* Selingo, *id.*, at 32-33.

12.   Only weeks after the Center for Individual Rights (CIR), architects of lawsuits against the University of Texas School of Law, the University of Michigan and its law school, and the University of Washington School of Law, placed full-page adds titled "Guilty by Admission" in fifteen college newspapers, the University of Massachusetts announced that it would dramatically curtail the use of race in admissions in order to protect itself from litigation. *See UMass Retreats From Race-Based*

their affiliated law schools have terminated race-sensitive affirmative action.[13] Furthermore, the Supreme Court will probably review one of the high-profile affirmative action cases currently pending against the University of Michigan,[14] the University of Michigan Law School,[15] the University of Georgia,[16] and the University of Washington Law School.[17] In December 2000, the Ninth Circuit upheld a lower court ruling in the *Smith* case that it is constitutionally permissible to use race as a plus factor in order to achieve educational diversity.[18] Consequently, there is a heightened possibility of Supreme Court review because of conflict between the Fifth Circuit (*Hopwood*) and the Ninth Circuit (*Smith*) over whether achieving a diverse student body is a compelling governmental interest sufficient to withstand strict scrutiny review.[19]

Affirmative Action, 27 J. BLACKS HIGHER EDUC. 12, 12 (2000); Mary Carey, *Policy or Practice?*, DAILY HAMPSHIRE GAZETTE, Mar. 26, 1999, at A1, *available at* 1999 WL 11723625; Mark Mueller, *UMass to Change Race-Based Policies*, BOSTON HERALD, Feb. 20, 1999, at 5, *available at* 1999 WL 3390642. CIR's advertisements declared, "Nearly Every Elite College in America Violates the Law. Does Yours?" *See* Theodore Cross, *African-American Opportunities in Higher Education: What Are the Racial Goals of The Center for Individual Rights?*, 23 J. BLACKS HIGHER EDUC. 94, 95-96 (1999); Laura Flanders, *Affirmative Racism*, THE NATION, Mar. 8, 1999, *available at* http://www.thenation.com/issue/990308/0308flanders.shtml. CIR provided free handbooks detailing strategies for students to obtain information on their universities and how to initiate lawsuits. Cross, *id.*

Like UMass, the University of Virginia's Board of Visitors directed the president to discontinue race-conscious admissions soon after CIR's accusatory advertisements ran in Virginia's *Cavalier Daily* in January 1999. *See* Peter Schmidt, *U. of Virginia Halts Use of Scoring System That Helped Black Applicants*, CHRON. HIGHER EDUC., Oct. 22, 1999, at A42. One difference is that although Virginia's admissions policy was substantially altered, university officials only quietly announced the change months later. *See id.*

13.   There are also important affirmative action cases involving highly competitive high schools, though that is not the focus of this Comment. *See, e.g.*, Wessman v. Gittens, 160 F.3d 790 (1st Cir. 1998) (limiting the use of race-conscious affirmative action under the diversity rationale beyond a narrow threshold); Preston Green, *May Examination Schools Use Racial Preferences in Their Admissions Process?:* Wessman v. Gittens, 135 EDUC. L. REP. 873 (1999) (analyzing the Wessman case).

14.   Gratz v. Bollinger, 183 F.R.D. 209 (E.D. Mich. 1998), *rev'd and remanded by* Grutter v. Bollinger, 188 F.3d 394 (6th Cir. 1999) (consolidated cases).

15.   Grutter v. Bollinger, 16 F. Supp. 2d 797 (E.D. Mich. 1998), *rev'd and remanded by* Grutter v. Bollinger, 188 F.3d 394 (consolidated cases).

16.   Johnson v. Bd. of Regents of the Univ. Sys. of Georgia, 106 F. Supp. 2d 1362 (S.D. Ga. 2000) (holding that race-sensitive affirmative action violates Title VI). The University of Georgia has since appealed this decision to the Eleventh Circuit. Sara Hebel, *U. of Georgia Plans Appeal of Affirmative-Action Ruling*, CHRON. HIGHER EDUC. Sept. 22, 2000, at A31.

17.   Smith v. Univ. of Wash. Law Sch., 2 F. Supp. 2d 1324 (W.D. Wash. 1998). The plaintiffs' and defendants' briefs are reprinted in 26 J.C. & U.L. 429 (2000) and 26 J.C. & U.L. 445 (2000), respectively.

18.   Smith v. Univ. of Wash. Law Sch., 233 F.3d 1188 (9th Cir. 2000); *see also* Sara Hebel, *U.S. Appeals Court Upholds Use of Affirmative Action in Admissions*, CHRON. HIGHER EDUC., Dec. 15, 2000, at A40.

19.   *See supra* note 7 for a discussion of *Hopwood*; *see also* Marcia Coyle, *It's Baa-aa-kke! A circuit split aims affirmative action at the Supreme Court*, NAT'L L.J., Jan 22, 2001, at A1 (predicting that the Supreme Court will likely reconsider *Bakke*); Sara Hebel, *Courting a Place in Legal History*, CHRON. HIGHER EDUC., Nov. 24, 2000, at A23 (reviewing several affirmative action cases and concluding that the Supreme Court is likely to review one of them).

A key issue in the litigation and policy clash over affirmative action is the role that standardized tests, like the SAT and the LSAT, play in the distribution of opportunity to attend highly selective institutions of higher education. For example, *Grutter v. Bollinger*, a challenge to affirmative action at the University of Michigan Law School, went to trial in January 2001 on three issues, one of which was "[w]hether race should be considered to offset disadvantages minority students face in test scores and grades."[20] Likewise, in *The Black-White Test Score Gap*, a recent book that analyzes the causes and consequences of racial differences in test performance, editors Christopher Jencks of Harvard and Meredith Phillips of UCLA argue that reducing the gap in standardized test scores is likely the most promising means of promoting greater educational and economic equality. Jencks and Phillips contend:

> [I]f racial equality is America's goal, reducing the black-white test score gap would probably do more to promote this goal than any other strategy that commands broad political support. Reducing the test score gap is probably both necessary and sufficient for substantially reducing racial inequality in educational attainment and earnings. Changes in education and earnings would in turn help reduce racial differences in crime, health, and family structure, although we do not know how large these effects would be.[21]

In this Comment, I attempt to inform the debate over affirmative action, currently playing out in the courts and academia, by vigorously investigating whether the LSAT may be biased against students of color. Subsequently, I will identify pitfalls in conventional thinking about "merit." I will also assess Jencks and Phillips's policy recommendation that the test score gap should be placed at the forefront of America's social agenda.

B. *The Adverse Impact of Standardized Testing in the Current Admissions Climate*

One strategy universities have recently adopted to promote racial and ethnic diversity in lieu of affirmative action is to de-emphasize standardized tests as criteria for entry. For example, the "One Florida" plan guarantees admission to at least one university in the state, though not necessarily to a student's top choice, to the top twenty percent of Florida's high school graduates, regardless of standardized test scores or other

---

20. Jodi S. Cohen, *Minorities Set to Testify at U-M Trial, Students Say Criteria Used for Law School Entry Discriminate,* DETROIT NEWS, Dec. 24, 2000, *available at* 2000 WL 30259961.

21. Christopher Jencks & Meredith Phillips, *The Black-White Test Score Gap: An Introduction,* *in* THE BLACK-WHITE TEST SCORE GAP 1, 3-4 (Christopher Jencks & Meredith Philips eds., 1998).

factors, such as school quality.[22] Texas and California have adopted similar programs. There is a "Ten Percent Plan"[23] in Texas and a "Four Percent Plan"[24] in California. Whereas the Texas plan guarantees admission to the most selective institutions in the system (UT-Austin and Texas A&M),[25] the California plan does not insure admission to the most competitive campuses of the UC system (Berkeley and UCLA).[26]

One commonality underlying these various "*X*-Percent Plans" is an implicit recognition that tests like the SAT and American College Testing (ACT) disproportionately exclude minority applicants in highly selective admissions. National data confirm that testing imposes a greater barrier than do other measures of performance. For example, Brookings Institution scholar William Dickens and Harvard Professor Thomas Kane recently analyzed the High School and Beyond (HSB) survey, a nationally representative sample of youth.[27] Dickens and Kane report that the HSB reveals a smaller Black-White gap in high school grades than in SAT scores.[28] Thus, in light of the ban on affirmative action, the University of California's Latino Eligibility Taskforce, chaired by Eugene Garcia, Dean of UC Berkeley's Graduate School of Education, recommended eliminating the SAT as a determinant of UC eligibility.[29] It is not surprising, then, that the prohibition on affirmative action accelerated already existing criticism of standardized testing.[30]

---

22. For a description of the "One Florida" plan, see Selingo, *supra* note 11 at 33; *Why the "One Florida" Plan Would Remove Blacks from the Best Campuses of the University of Florida, supra* note 11 at 29-30.

23. For an analysis of the Texas "Ten Percent Plan," see Forbath & Torres, *supra* note 7, at 185; Danielle Holley & Delia Spencer, Note, *The Texas Ten Percent Plan*, 34 HARV. C.R.-C.L. L. REV. 245, 262-77 (1999).

24. The UC Regents approved the "Four Percent Plan" in March 1999. *See* Pamela Burdman, *UC Regents Rethinking Use of SAT*, S.F. CHRON., Mar. 20, 1999, at A22; V. Dion Haynes, *U of California Alters Its Policy on Admissions*, CHI. TRIB., Mar. 20, 1999, *available at* 1999 WL 2855179. The "Four Percent Plan" is a scaled-down version of a constitutional amendment proposed in 1997 by state Senator Teresa Hughes (D-Inglewood) that would have offered admission to the top 12.5% of California's high school graduates. *See* Kenneth R. Weiss, *Plans Seek More UC Pupils From Poorer Schools*, L.A. TIMES, May 12, 1997, at A1.

25. Holley & Spencer, *supra* note 23, at 246.

26. Burdman, *supra* note 24, at A22.

27. William T. Dickens & Thomas J. Kane, *Racial Test Score Differences as Evidence of Reverse Discrimination: Less than Meets the Eye*, 38 INDUS. REL. 331, 361-62 (1999) (describing the authors' use of the HSB databases).

28. *Id.* at 338. Using the National Longitudinal Survey of Youth (NLSY), Dickens and Kane similarly found a smaller Black-White gap in schooling than on the Armed Forces Qualifying Test. *Id.*

29. *See* UNIV. OF CAL. LATINO ELIGIBILITY TASKFORCE, LATINO STUDENT ELIGIBILITY AND PARTICIPATION IN THE UNIVERSITY OF CALIFORNIA: YA BASTA!, Report No. 5, 19 (1997); *see also* Z. Byron Wolf, *Task Force Urges Regents to Drop SAT Requirement*, DAILY CALIFORNIAN, Sept. 19, 1997, at 1.

30. *See* Edward G. Haggerty, *LSAT: Uses and Misuses*, N.Y. BAR J., May/June 1998, at 45. Haggerty, Director of Media Relations for LSAC, reports, "With affirmative action in university admissions under attack on many fronts, public scrutiny of standardized admission tests has perhaps never been greater. At the same time, reliance on tests like the LSAT in contexts for which they are not

CALIFORNIA LAW REVIEW   [Vol. 89:1055

Criticism of the testing gained further momentum when, in February 2001, UC President Richard Atkinson also recommended that the UC system discontinue using the SAT I in the admissions process.[31] In a widely reported[32] speech to the American Council on Education, Atkinson concluded, "America's overemphasis on the SAT is compromising our educational system."[33] Atkinson's recommendation will likely reverberate nationally, especially with respect to the unfolding affirmative action debate.[34]

While the *X*-Percent Plans are at least an option at the undergraduate level, they are nonexistent in law school admissions. The U.S. Commission on Civil Rights recently concluded that a "major problem with the percentage plans is their inattention to law schools, medical schools, and other graduate and professional schools, where ending affirmative action is devastating."[35] For example, at UC law schools in the four years since Proposition 209 and SP-1 took effect, African American enrollments dropped from 7.5% to 2.2% compared to the four previous years, Chicano/Latino enrollments dropped from 13.4% to 7.2%, and Native American enrollments went from 1.4% to 0.7%.[36]

Soon after the implementation of California's affirmative action ban, several civil rights organizations filed a complaint with the Office for Civil Rights (OCR) of the U.S. Department of Education regarding the admission policies at University of California law schools.[37] They alleged,

---

designed or validated seems to be increasing." *Id.*; Pamela Burdman, *Worth of SAT Exam Questioned*, S.F. CHRON., Nov. 11, 1997, at A1; Bryan Mealer, *Moves Against Affirmative Action Fuel Opposition to Standardized Admissions Tests*, CHRON. HIGHER EDUC., Oct. 17, 1997, at A40.

31. Jeffrey Brainard, *U. of California's President Proposes Dropping the SAT Requirement*, CHRON. HIGHER EDUC., Feb. 19, 2001, at daily online version http://chronicle.com/cgi2-bin/printable.

32. *See, e.g.*, Diana Jean Schemo, *Head of U. of California Seeks to End SAT Use in Admissions*, N.Y. TIMES, Feb. 17, 2001, at A1; Kenneth R. Weiss, *SAT May Be Dropped as UC Entrance Exam*, L.A. TIMES, Feb. 17, 2001 at A1; *Nightline: The Big Test* (ABC television broadcast, Mar. 15, 2001); John Cloud, *Should SATs Matter?*, TIME, Mar. 4, 2001 at 41.

33. Richard C. Atkinson, Robert H. Atwell Distinguished Lecture at the American Council on Education Annual Meeting (Feb. 18, 2001) (transcript available at http://www.ucop.edu/ucophome/commserv/sat/speech.html).

34. Jeffrey Selingo & Jeffrey Brainard, *Call to Eliminate SAT Requirement May Reshape Debate on Affirmative Action*, CHRON. HIGHER EDUC., Mar. 2, 2001, at A21.

35. U.S. COMM'N ON CIVIL RIGHTS, TOWARD AN UNDERSTANDING OF PERCENTAGE PLANS IN HIGHER EDUCATION: ARE THEY EFFECTIVE SUBSTITUTES FOR AFFIRMATIVE ACTION? 5 (2000). Many of the opinions expressed in this report are summarized in Mary Frances Berry, *How Percentage Plans Keep Minority Students Out of College*, CHRON. HIGHER EDUC., Aug. 4, 2000, at A48; *see also* Russell, *supra* note 7, at 508 ("At the law school, however, the 10% plan has no application, and although the law school's admissions committee has recrafted the criteria for admission, African American and Latino law students have not returned to pre-*Hopwood* levels.").

36. William C. Kidder, *Affirmative Action in Higher Education: Recent Developments in Litigation, Admissions, and Diversity Research*, 12 LA RAZA L.J. (forthcoming 2001), *available at* http://www.scu.edu/law/salt/affirmative/affirmative.html (visited Mar. 14, 2001).

37. Peter Schmidt, *New Admissions Policy at U. of California Graduate Schools Called Discriminatory*, CHRON. HIGHER EDUC., Mar. 28, 1997, at A43.

in part, that by relying so heavily on the LSAT, the UC law schools' current admission policies have a disparate impact on underrepresented minority applicants, in violation of Title VI of the 1964 Civil Rights Act.[38] In response, OCR launched an ongoing compliance review of the law schools at UC Berkeley, UCLA, and UC Davis.[39]

Others have recognized the racial gatekeeping effect of the LSAT. An American Bar Association (ABA) Committee on Diversity in Legal Education recommended using other selection criteria once a qualifying threshold has been met.[40] Also, a December 1999 Law School Admission Council ("LSAC") admission policy statement reported that pressure from threatened affirmative action litigation and law school rankings were promoting misuse of the LSAT in law school admissions.[41] The report concluded that the vast majority of ABA-accredited law schools may be using the LSAT inappropriately.[42]

In *Grutter v. Bollinger*,[43] a challenge to the affirmative action program at the University of Michigan Law School, the parties also disputed the fairness of the LSAT. The University employed Stanford psychologist Claude Steele, whose experimental research on "stereotype threat" has intensified the debate over test bias, as an expert witness on testing.[44] Student of color organizations, intervening as defendants, submitted reports from several expert witnesses on the issue of bias in standardized testing. One expert, David M. White, devoted a substantial portion of his report to establishing and explaining the discriminatory impact of the LSAT vis-à-vis undergraduate grades for African American, Latino, and Native American applicants.[45]

---

38. Stephen Burd, *Education Dept. Delays Loan Consolidations*, CHRON. HIGHER EDUC., July 25, 1997, at A32.

39. *Id.*; Louis Freedberg, *UC's Law Admissions Investigated*, S.F. CHRON., July 15, 1997, at A1; Amy Wallace, *UC Law Schools Face Discrimination Investigation*, L.A. TIMES, July 15, 1997, at A3.

40. *See* A.B.A., REPORT OF THE COMMITTEE ON DIVERSITY IN LEGAL EDUC. *passim* (1998). Many of the Committee's suggestions are discussed in Michael A. Olivas, *Higher Educ. Admissions and the Search for One Important Thing*, 21 U. ARK. LITTLE ROCK L. REV. 993, 1007-11 (1999).

41. *See* LAW SCH. ADMISSION COUNCIL, NEW MODELS TO ASSURE DIVERSITY, FAIRNESS, AND APPROPRIATE TEST USE IN LAW SCH. ADMISSIONS 4, 7, 10 (1999) [hereinafter NEW MODELS TO ASSURE DIVERSITY]. For a short summary, see *Notebook*, CHRON. HIGHER EDUC., Jan. 7, 2000, at A59.

42. *See* NEW MODELS TO ASSURE DIVERSITY, *supra* note 41, at 21.

> Many, perhaps more than 90 percent of law schools, currently use what is commonly known as a presumptive admission model. This method arrays candidates on the basis of their index scores and identifies a group at the top that can be approved for admission strictly on the basis of the index score, and a group at the bottom slated for denial .... Schools currently using the model described above are encouraged to modify it because such methods may be using the LSAT score incorrectly.

Id.

43. 16 F. Supp. 2d 797 (E.D. Mich. 1998).

44. *See infra* Part.IV.B.

45. Expert Report of David M. White, Grutter v. Bollinger, 16 F. Supp. 2d 797 (E.D. Mich. 1998) (No. 97-75928). This and the other expert reports for the intervening defendants will be reprinted in the Fall 2001 issue of the *La Raza Law Journal*.

Furthermore, Testing for the Public's study of 1993-94 to 1997-98 admission decisions at over 175 ABA-accredited law schools demonstrated the adverse consequences for students of color stemming from over-reliance on the LSAT.[46] This study documented that Whites had the highest admission rates of any racial or ethnic group among applicants with equivalent UGPAs, even with affirmative action.[47] Overall, 72% of White applicants were admitted to at least one law school, compared to 46% for African Americans, 60% for Hispanics, 61% for Chicanos, 69% for Asian Americans, and 62% for Native Americans.[48] Needless to say, the consequences have been far more severe for students of color at law schools where it is now impermissible to consider race or ethnicity in admission decisions.[49]

In summary, the available data from undergraduate institutions and law schools suggest that standardized tests typically produce larger differences between Whites and students of color than other academic criteria. This fact negatively affects minorities' admission opportunities, particularly when test scores are heavily relied upon or when race-conscious affirmative action is prohibited.

C. Responses to Criticism of Standardized Testing

While some affirmative action proponents have criticized the greater discriminatory impact of standardized tests on underrepresented minority students, these criticisms have not gone unchallenged. Harvard Professor Stephan Thernstrom and Manhattan Institute Fellow Abigail Thernstrom, leading opponents of race-sensitive affirmative action, counter that the SAT gap between White and Black students is a neutral reflection of differences in academic preparation, and, as such, is no larger than the gap in other measures of achievement. Responding to testing critics in their book *America in Black and White*, the Thernstroms argue:

> When they heap scorn on "mere tests," defenders of affirmative action pick an easy target, and deflect attention away from a wealth of evidence demonstrating that the racial gap in other measures of

---

46.   For a short summary of these findings, see Katherine S. Mangan, *Study Says Law Schools Favor White Applicants over Their Minority Peers with the Same Grades*, CHRON. HIGHER EDUC., June 27, 2000, *available at* http://www.thechronicle.com.

47.   *See* William C. Kidder, *Portia Denied: Unmasking Gender Bias on the LSAT and Its Relationship to Racial Diversity in Legal Education*, 12 YALE J.L. & FEMINISM 1, 13-14 (2000) [hereinafter Kidder, *Portia Denied*]; William C. Kidder, *The Rise of the Testocracy: An Essay on the LSAT, Conventional Wisdom, and the Dismantling of Diversity*, 9 TEX. J. WOMEN & L. 167, 209 tbl.3 (2000) [hereinafter Kidder, *Rise of the Testocracy*].

48.   Kidder, *Portia Denied*, *supra* note 47, at 14 tbl.4; Kidder, *Rise of the Testocracy*, *supra* note 47, at 209 tbl.3.

49.   *See infra* notes 201-202 for discussion of admission to public law schools in Texas. For the consequences of Prop. 209 at UC law schools, see Kidder, *Situating Asian Pacific Americans*, *supra* note 9, at 53-45.

academic achievement and preparation is just as large as the gap in SAT scores. . . . So far, at least, critics of tests have been unable to demonstrate that any other measure of academic preparation and achievement yields a significantly different result.[50]

UC Santa Barbara Professor Rebecca Zwick, formerly of the Educational Testing Service (ETS), also argues that the magnitude of racial and ethnic gaps on the SAT is substantially equivalent to the gaps in high school grades.[51] Furthermore, a widely cited UC Office of the President study questioned the recommendations of the UC Latino Eligibility Taskforce, and argued that if the UC system eliminated the SAT, and raised the high school GPA requirement correspondingly to hold constant the total number of UC-eligible students, eligibility rates would remain the same for Latinos and would drop for African Americans.[52]

In the law school arena, LSAC officials similarly emphasize that the difference between admitting students by LSAT scores or UGPAs is inconsequential.[53] Some argue that the equivalent size of the racial or ethnic gaps on LSAT scores and college grades can be misleading because underrepresented minority students disproportionately enroll in less rigorous majors, or in less competitive colleges, or both. Professor David Kaye,

---

50. STEPHAN THERNSTROM & ABIGAIL THERNSTROM, AMERICA IN BLACK AND WHITE: ONE NATION, INDIVISIBLE 402-03 (1997). The Thernstroms reviewed evidence on class rank, grade-point averages, and course selection in making their claim about the comparable gaps in academic achievement. *Id.* For a critique of the conclusions the Thernstroms draw from this data, see Stephen R. Shalom, *Dubious Data: The Thernstroms on Race in America*, 1 RACE & SOC'Y 125, 132-33 (1998). For a critique of Stephan Thernstrom's analysis of law school admissions after Proposition 209, see Kidder, *Situating Asian Pacific Americans, supra* note 9, *passim.*

51. Rebecca Zwick, *Eliminating Standardized Tests in College Admissions: The New Affirmative Action?*, 81 PHI DELTA KAPPAN, 320, 323 (1999) ("Because the pattern of ethnic group differences in average high school GPA is usually similar to the pattern of average admissions test scores, an admissions policy that excludes tests but continues to include high school grades is unlikely to produce dramatic change." Zwick concludes: "In any case, both test scores and high school grades are reflections of the very same disparities in educational opportunity."); *see also* Arthur R. Jensen, *Testing: The Dilemma of Group Differences*, 6 PSYCHOL., PUB. POL'Y & L. 121, 123 (2000) ("[T]he sole use of GPA for selection usually results in a highly similar ranking of applicants, and strict top-down selection still has almost as much adverse impact as test scores or even test scores and GPA combined.").

52. UNIV. OF CAL. OFFICE OF THE PRES., UNIVERSITY OF CALIFORNIA FOLLOW-UP ANALYSES OF THE 1996 CPEC ELIGIBILITY STUDY 25 (1997). This report (or at least its conclusion) is widely cited by defenders of standardized testing. *See* June Kronholz, *As States End Racial Preferences, Pressure Rises to Drop SAT to Maintain Minority Enrollment*, WALL ST. J., Feb. 12, 1998, at A24 (noting that the College Board, which administers the SAT, rebuts the UC Latino Eligibility recommendation by arguing that eliminating the SAT would cause the White and Asian eligible pools to increase even more); David W. Murray, *The War against Testing*, COMMENTARY, Sept. 1998, at 34; Shannon Starr, *Campuses Find Diversity Is Harder to Maintain*, PRESS-ENTERPRISE (Riverside, Cal.), Feb. 20, 2000, at B12 (quoting Keith Widaman, a UC Davis professor and member of the UC Board of Admissions: "Dropping the SAT sounds like it would result in more black and Latino students gaining admission. But when you look at the whole picture, whites and Asian numbers would increase proportionately, as well."); Zwick, *supra* note 51, at 323.

53. These claims are assessed in greater detail *infra* Part IV.C.2.

an expert on the use of statistics in litigation, and a long-time member of LSAC's Test Development and Research Committee, predicted that with proper controls for college performance, the LSAT might not produce a racial screening effect at all. Comparing LSAT scores and UGPAs, Kaye claims "if attention were paid to such factors as the field of study and the college attended, the gains associated with undergraduate grades might disappear."[54]

This study is an effort to provide empirical answers to the ongoing scholarly debate over whether the LSAT stratifies opportunity by race and ethnicity among students who have demonstrated similar accomplishment levels in college. To reframe the question in the Thernstroms's terms, the present investigation seeks to answer whether "any other measure of academic preparation and achievement yields a significantly different result"[55] than that obtained on the LSAT. This question will be addressed in the manner suggested by Professor Kaye. This study will compare LSAT performance among students with equivalent UGPAs, and control for factors such as institution attended, graduation date, and choice of major.

II
METHODOLOGY

A.   Characteristics of the Data Sample

Boalt Hall is a highly competitive public law school that draws a national pool of applicants similar to other "elite" law schools across the country. In recent years, Boalt Hall received between 4000 and 6000 annual applications for an entering class of 270, with the average applicant possessing a 3.4 UGPA and a 161 on the LSAT (86th percentile).[56] The 1996, 1997, and 1998 applicant pools to Boalt Hall provided the data for this study.[57] For each applicant, anonymous data was obtained on race and ethnicity, undergraduate institution attended, graduation date, cumulative UGPA (reported to the nearest hundredth of a point), age (in days), and LSAT score (120-180 scale).[58]

After the applicant pools from the three years were collapsed into a single database, applicant data was categorized by undergraduate

---

54.   David Kaye, *Searching for Truth about Testing*, 90 YALE L.J. 431, 456 n.100 (1980) (book review).

55.   THERNSTROM & THERNSTROM, *supra* note 50, at 403.

56.   UNIV. OF CAL., BERKELEY SCH. OF LAW (BOALT HALL), ANNUAL ADMISSIONS REPORT 3 (1997).

57.   More precisely, this means the candidates from the 1995-96, 1996-97, and 1997-98 admission cycles, or those who applied for Boalt's 1996, 1997, and 1998 first-year entering classes.

58.   LSAT scores for applicants who took the test more than once were reported as an average score, which is the way LSAC traditionally recommends treating multiple scores in the absence of extenuating circumstances such as illness. SUSAN P. DALESSANDRO & LORI D. MCLEOD, LAW SCH. ADMISSION COUNCIL, THE VALIDITY OF LAW SCHOOL ADMISSION TEST SCORES FOR REPEATERS: A REPLICATION 13 (1999).

institution.[59] Further analysis was conducted on applicants from fifteen colleges and universities, four public and eleven private, that were considered to be top "feeder schools" to Boalt Hall based on applicant volume, and also were recognized as "highly selective" institutions.[60] The schools studied are listed in descending order by applicant volume: UC Berkeley, UCLA, Stanford, Harvard, Yale, Cornell, Brown, Princeton, University of Michigan-Ann Arbor, UT Austin, Northwestern, Dartmouth, Georgetown, Duke, and Columbia. Thus, the schools studied reflect a large geographic distribution. The above fifteen institutions alone account for 40% of Boalt's 1996-98 applicants and 53% of its admission offers.[61]

These undergraduate schools not only provide a large number of applicants to Boalt, but to the nation, since applicants from six of the top ten (and each of the top four) providers of U.S. law school candidates are included in this study.[62] Seven of the fifteen schools in this study are also universities in the Andrew W. Mellon Foundation's "College and Beyond" database, used by William Bowen and Derek Bok in *The Shape of the*

---

59. Collapsing three separate pools means that a small number of persistent applicants who were denied one year and reapplied the next are counted twice. It was decided that the benefits of combining several years of data, making possible matches between a much higher number of students, far outweighed this modest drawback of double-counting. No information is available about differential reapplication rates by race or ethnicity, or rates of reapplication generally. However, nationally during 1996-98 the average applicant applied to over four ABA-accredited law schools, and received at least one offer of admission 69% of the time. LAW SCH. ADMISSION COUNCIL, 1997-98 NATIONAL DECISION PROFILES (1999); LAW SCH. ADMISSION COUNCIL, 1996-97 NATIONAL DECISION PROFILES (1998); LAW SCH. ADMISSION COUNCIL, 1995-96 NATIONAL DECISION PROFILES (1997). Among those admitted, the yield rate (the proportion of students who chose to enroll) for these three cycles was 82%. *See id.* Applicants who reapplied either received no admission offers or were dissatisfied with the offer(s) they received and were sufficiently motivated to go through the application process all over again. For the fifteen "elite" universities in the database, I suspect that cumulative acceptance rates are significantly higher. All this is to say that the impact of double counting reapplicants most likely has a negligible influence in terms of creating a sampling bias.

60. The choice of highly selective universities is not to suggest that applicants from other institutions do not or cannot make outstanding law school candidates. Rather, the goal was to define a fairly homogenous group of schools in order that generalizations could be extended beyond Boalt's specific applicant pool. Given the characteristics of Boalt Hall's applicant pool, it made the most sense to organize the study around "elite" institutions. This was particularly so since UC Berkeley, UCLA, Stanford, Harvard, and Yale, respectively, produced the most applicants to Boalt during the 1996-98 period. With a different database, a different emphasis would have been appropriate. Of additional significance is that the adverse consequences of eliminating affirmative action have shown to be most severe at highly selective schools like UC Berkeley, making this set of schools of heightened relevance to policy makers. *University of California: Black Students Moving Down into the Less Selective Campuses,* 19 J. BLACKS HIGHER EDUC., 28, 28 (1998).

61. *See* UNIV. OF CAL., BERKELEY SCH. OF LAW (BOALT HALL), ANNUAL ADMISSIONS REPORT tbl. V (1998); UNIV. OF CAL., BERKELEY SCH. OF LAW (BOALT HALL), ANNUAL ADMISSIONS REPORT tbl. V (1997); UNIV. OF CAL., BERKELEY SCH. OF LAW (BOALT HALL), ANNUAL ADMISSIONS REPORT tbl. V (1996).

62. *The Leading Undergraduate "Feeder Schools" for Black and White Law School Applicants,* 26 J. BLACKS HIGHER EDUC. 85, 85 (Winter 1999/2000) (listing the top feeder schools in 1997-98 in terms of overall applicant volume, including UCLA (#1), UC Berkeley (#2), UT (#3), University of Michigan (#4), Cornell (#8), and Harvard (#9)).

River,[63] a mammoth study of long-range outcomes for African American and White graduates of highly selective colleges and universities. I believe the database contains a sufficiently large fraction of the highly selective law school "feeder institutions" that the findings of this study can be applied to law school applicants from elite American universities with reasonable confidence.[64]

The data sample also excludes small numbers of applicants. For example, the sample does not include those who are missing data such as an LSAT score or UGPA.[65] Puerto Rican applicants were excluded due to extremely small sample size. Applicants who listed their ethnicity as "foreign" or "other" were excluded, as were those who "declined to state" their ethnicity.[66] While some scholars speculate that declining enrollment for underrepresented minorities at UC has been exaggerated due to this concomitant "decline to state" trend,[67] available evidence suggests otherwise, since "decline to state" applicants have been found to be overwhelmingly White.[68]

Still other applicants were excluded because they graduated prior to 1993.[69] For purposes of this matching study, I concluded that a fairly homogenous group of 1993 to 1998 graduates would largely control for variations of grading standards within institutions over time. This issue is

---

63. WILLIAM G. BOWEN & DEREK BOK, THE SHAPE OF THE RIVER: LONG-TERM CONSEQUENCES OF CONSIDERING RACE IN COLLEGE AND UNIVERSITY ADMISSIONS 40 (1998). The universities in both the "College and Beyond" and the present databases are Stanford, Columbia, Duke, Princeton, Michigan, Northwestern, and Yale. The "College and Beyond" database is more heavily weighted with small liberal arts colleges (coed and women's) such as Williams and Wellesley.

64. While the present results do not include a broad enough range of institutions to infer that the findings can be automatically applied to law school applicants generally, it is noteworthy that Gannon's study included a more nationally representative cross section of law school applicants and his results were remarkably similar to those reported herein. Gannon, *supra* note 5, at 274-75, 277.

65. It is likely that many of these applicants did not follow through with the entire application process.

66. After SP-1 and Proposition 209, "decline to state" applications increased significantly at both the law schools and the undergraduate campuses of the University of California. *See* Kidder, *Situating Asian Pacific Americans*, *supra* note 9, at 44 n.72 (reviewing trends in "decline to state" applications at the University of California); *see also Fewer U. of California Applicants Reveal Race*, N.Y. TIMES, Apr. 2, 1998, at A4.

67. *See* Martin Trow, *California after Racial Preferenees*, 135 PUB. INTEREST 64, 69 (1999) (arguing that the published figures for minority enrollments at UC schools are "almost certainly lower than the true figure" due to the "decline to state" trend).

68. *See* Kidder, *Situating Asian Pacific Americans*, *supra* note 9, at 44 n.72 (reviewing evidence from Boalt Hall, UCLA Law School, and the overall UC system at the undergraduate level). In order for "decline to state" applicants to pose a serious problem of sampling bias, it would need to be true that "decline to state" White applicants differ significantly from other White applicants on the LSAT (or another factor that significantly correlates with the LSAT). However, LSAT averages for White and "decline to state" applicants are nearly identical for the entire Boalt Hall applicant pool.

69. In a few rare instances where graduation date was the only missing information, age was used as a proxy for graduation date, and applicants over 9000 days old (24.6 years) were tossed out. This decision was made because, while it cannot be assumed that older students would have graduated prior to 1993, it is safer to assume that younger applicants must have graduated after 1992.

of particular concern since evidence indicates that the pattern of "grade inflation" over the years has been greater at elite colleges like the ones in this database.[70]

B.   Method of Matching Applicants and Computing Basic Comparisons

The procedure for matching minority and White applicants with equivalent UGPAs from the same institution is modeled after the Gannon study discussed in Part IV.A. The primary objective of this study was to examine possible racial and ethnic differences in LSAT performance among applicants who demonstrated a comparable level of academic achievement over their undergraduate careers. Applicants within each college were separated into five racial categories, consistent with general Department of Education and university classifications: White/Caucasian, Black/African American, Chicano/Latino, Native American/American Indian, and Asian Pacific American.[71]

Since the Asian Pacific American category is quite heterogeneous, it is fortunate that Boalt's top three feeder schools, UC Berkeley, UCLA, and Stanford, are also three of the highly selective universities with the greatest percentage of Asian Pacific Americans in the country.[72] This permitted

---

70.   *See, e.g.,* Randal C. Archibold, *At Princeton, the 'Gentlemen's C' Would Not Seem So Gentlemenly,* N.Y. TIMES, Feb. 18, 1998, at A1; Ben Gose, *Duke May Shift Grading System to Reward Students Who Take Challenging Classes,* CHRON. HIGHER EDUC., Feb. 14, 1997, at A40; Ben Gose, *Efforts to Curb Grade Inflation Get an F From Many Critics,* CHRON. HIGHER EDUC., July 25, 1997, at A41; Noel Perrin, *How Students at Dartmouth Came to Deserve Better Grades,* CHRON. HIGHER EDUC., Oct. 9, 1998, at A68. In contrast to highly selective colleges and universities, analysis by the U.S. Department of Education suggests that there has not been grade inflation nationally in recent decades. *See* CLIFFORD ADELMAN, U.S. DEP'T OF EDUC., THE NEW COLLEGE COURSE MAP AND TRANSCRIPT FILES 198, 200 (2nd ed. 1999) (reviewing 21,000 transcripts at 3000 colleges and finding no grade inflation when comparing 1972-82 with 1982-92); *see also* Clifford Adelman, *A's Aren't That Easy,* N.Y. TIMES, May 17, 1995, at A19. Earlier research based on national data provided by LSAC as well as colleges and universities indicates that the highest (that is, the most inflated) grade distributions occur at highly selective institutions, whereas the lowest grade distributions occur at historically Black institutions. *See* Sandra W. Weckesser, *The Double Jeopardy of the GPA: A Comparative Study of Grade Distribution Patterns and Grade Inflation by Types of Colleges and Universities, in* TOWARDS A DIVERSIFIED LEGAL PROFESSION, *supra* note 5, at 284, 296-97 tbl.1.

71.   *See, e.g.,* UC Office of the President, Law School Applicants, Admits, and First Year Class Enrollments, *at* http://www.ucop.edu/adadadv/lawdata/lawschl2.html (last modified Sept. 6, 2000) (categorizing law school applicants by these five classifications).

72.   This is particularly true at UC Berkeley and UCLA, where Asian Pacific Americans have outnumbered Whites for several years. *See* UNIV. OF CAL., BERKELEY OFFICE OF STUDENT RESEARCH, BERKELEY UNDERGRADUATE FACT SHEET-FALL 1998 (1999) (listing Asian Americans as 40.9% of UC Berkeley undergraduates, compared to 31.3% for Whites); *see also* Elaine H. Kim, *Visible But Marginal: Changing the Status of Asian Pacific Americans at UC Berkeley,* 6 UCLA ASIAN PAC. AM. L.J. 55, 58 (2000) (noting that Asian Americans make up about 40% of UC Berkeley undergraduates, compared to 35% for Whites). Furthermore, in 1996-97, Asian Pacific Americans comprised 24% of Stanford's undergraduates. Stephan Thernstrom & Abigail Thernstrom, *Reflections on The Shape of the River,* 46 UCLA L. REV. 1583, 1629 (1999). This pattern is partly a byproduct of the concentration of Asian Pacific Americans in California. *See* Kidder, *Situating Asian Pacific Americans, supra* note 9, at 37.

CALIFORNIA LAW REVIEW [Vol. 89:1055

separate reporting of LSAT performance for several Asian Pacific American subgroups, including Chinese, Japanese, Filipino, Korean, Southeast Asian, and East Indian or Pakistani. To the best of my knowledge, this is the first reporting of its kind; previous law school applicant information separated by Asian national or ethnic groups has been entirely unavailable in published LSAC research and the legal education literature.

Within each racial and ethnic group, applicants were matched according to academic equivalency. Academic equivalency was operationally defined as applicants who had the same cumulative UGPA within plus or minus .10 on a 4.0 scale.[73] As Gannon noted in his 1981 study:

> A bandwidth of plus or minus .10 is considered reasonably narrow for operationally defining equivalence in academic performance. After all, there is a certain standard error of measurement associated with college grades, and law school admissions officers express a reluctance to make practical, selection/rejection decisions based on such a slim margin of difference in applicants' UGPAs.[74]

In order to guard against the possibility that the +/-.10 bandwidth might artificially magnify racial LSAT differences, compared to what would occur if applicants were matched only to those with identical UGPAs[75] (to the nearest hundredth of a point), additional empirical tests were run on a subset of the data. On the one hand, using a +/-.10 bandwidth was found to have no such magnification effect compared to a zero bandwidth.[76] On the other hand, requiring identical UGPAs, instead of matching within a reasonable bandwidth, would have compromised the value of this study by

---

73. Technically speaking, it is slightly inaccurate to state that a 4.0 scale was used. The Law School Admission Council standardizes grades for all domestic ABA-accredited law school applicants. LSAC converts "A+" grades into a 4.33 instead of the 4.0 that "A's" are assigned. *See* LAW SCH. ADMISSION COUNCIL, LSAT REGISTRATION & INFORMATION BOOK 22 (2000) (describing the grade conversion process employed by the Law School Data Assembly Service). Thus an applicant with a perfect college record and a couple of "A+'s" can obtain a cumulative UGPA over 4.0. As a practical matter, this was the case for only a handful of the thousands of applicants in the study. Conversely, LSAC converts "No Pass" grades to a 0.0 ("F"), whereas many colleges exclude "No Pass" grades from UGPA calculations. Both practices tend to slightly unpack compressed UGPA distributions.

74. Gannon, *supra* note 5, at 276.

75. Hypothetically, the sampling bias I refer to could conceivably result from minorities having lower average UGPAs. The combination of the +/-.10 bandwidth and higher White UGPAs would, on the whole, cause minorities to be matched with Whites who had slightly higher UGPAs. An example clarifies the point. If a Black applicant from UC Berkeley with a 3.25 UGPA is compared with one hundred White UC Berkeley applicants with 3.15-3.35 UGPAs, the Black applicant actually would be matched with forty Whites with lower UGPAs, ten Whites who were exact matches, and fifty Whites with slightly higher UGPAs. On the other hand, this process of accumulating minute matching biases would start to reverse direction once matching was conducted above the White mean UGPA.

76. Matching candidates with identical UGPAs to the nearest hundredth was conducted for UC Berkeley applicants. Comparing the exact UGPA matching to the +/-.10 bandwidth matching for UC Berkeley applicants reported in Part III, average LSAT differences by race were all within the same one point range using either method. Exact matching caused the gap between African Americans and Whites to go up a bit and for the Chicano/Latino-White gap and the Asian Pacific American-White gap to go down modestly.

substantially lowering the total number of possible minorities who could be matched, and it would have even more dramatically curtailed the number of White applicants who could be matched with those minorities. Thus, the benefits of employing a +/-.10 bandwidth research design greatly outweighed the costs in comparison to the identical matching option.

After applicants in each cohort were ranked by UGPA, each minority applicant was matched with all the White candidates possessing UGPAs from the same school within +/-.10.[77] At this point, each minority's LSAT score was subtracted from the average LSAT for the White comparison group. For example, the LSAT score for each student of color with a 3.44 UGPA from UC Berkeley was subtracted from the LSAT average for the eighty-one White UC Berkeley applicants matched on UGPA (3.34-3.54). This process of computing LSAT score differences among UGPA-matched applicants continued for each and every matchable applicant in each institution. Overall LSAT differences were computed based on a weighted average; that is, they were weighted by the number of matches corresponding to each comparison, of all the recorded individual LSAT differences. In effect, this methodology was the same as matching every possible minority student in a cohort to every possible White student with equivalent academic performance from the same school.

The +/-.10 bandwidth boosted the size of the matching sample by permitting overlapping comparisons within the same applicant pool. Returning to our UC Berkeley example, each minority applicant with a 3.45 UGPA is also matched with eighty-one White applicants with 3.35-3.55 UGPAs. Most, but not all, of these White students were the same ones matched in the 3.34-3.54 group discussed earlier. Together, weighted averaging and multiple-matching reduces the influence of the unusual student on the overall results.

III

RESULTS

A. Basic Matching

The results of the matching, shown in Table 1, demonstrate a significant gap in LSAT scores between White students and students of color. The data reveals this result in spite of efforts to equalize the applicants being compared in terms of UGPA, graduation date, and institution attended. Among law school applicants with roughly equal educational attainment over their college years, the vanguard of high-achieving students of color,

---

77.   This meant that in a small number of cases (mostly involving unusually high or low UGPAs), a minority or White applicant was without a relevant comparison group. These unmatched applicants were eliminated from further study.

and African Americans in particular, pay a heavy "poll tax"[78] when the LSAT is included as an influential admission requirement.

TABLE 1:
AVERAGE LSAT SCORE DIFFERENCES BY RACE/ETHNICITY
STUDENTS WITH EQUIVALENT PERFORMANCE IN THE SAME
COLLEGE

Racial/Ethnic Group	African American	Chicano/ Latino	Native American[79]	Asian Pacific American
No. Matched	247	407	33	1043
LSAT Gap (White - Minority)	9.2	6.8	4.0	2.5

Even after implementing the controls, African Americans trail their equally accomplished White classmates by 9.2 points on the LSAT, with Chicanos and Latinos 6.8 points behind, Native Americans 4.0 points lower, and Asian Pacific Americans 2.5 points behind. To place these figures in perspective, the Black-White gap of 9.2 points is greater than a one standard deviation difference in the national applicant pool.[80]

---

78.   For discussion of standardized tests as modern day educational poll taxes for students of color and those from modest socioeconomic backgrounds, see Kidder, *Rise of the Testocracy, supra* note 47, at 209; Susan Sturm & Lani Guinier, *The Future of Affirmative Action: Reclaiming the Innovative Ideal*, 84 CALIF. L. REV. 953, 957 (1996); *LSAT: A Testing "Poll Tax" on Minority Law School Applicants*, FAIRTEST EXAMINER, Summer 2000, at 3.

79.   As I worked with the data, I was struck by how few Native Americans attended elite colleges and universities. Of the thirty-three Native American applicants matched on UGPA with Whites, twenty-six (79%) attended either UC Berkeley or UCLA. For an analysis of the undergraduate and law school applicant pools of Native Americans, see Gloria Valencia-Weber, *Law School Training of American Indians As Legal-Warriors*, 20 AM. INDIAN L. REV. 5, 27-36 (1995-96). Valencia-Weber reports that Oklahoma and California are consistently the two states that produce the highest number of Native American law school applicants. *See id.* at 34. She also reports that in the 1993-94 admission cycle, the top five producers of Native American law school applicants were UCLA (25); University of Oklahoma (24); Northeastern Oklahoma State University (21); UC Berkeley (16); and Stanford (10). *Id.* at 35 n.95. The ban on affirmative action correlated with a substantial drop in Native American enrollments at UC law schools since it was put into effect in 1997. At Boalt Hall, UC Davis School of Law, and UCLA School of Law there were fourteen combined first-year Native American enrollments in 1997-99 (0.63% of the class), compared to thirty-three first-year Native American enrollments in 1994-96 (1.53% of the class). *See* Memorandum from Dean Herma Hill Kay to the Boalt Community (Sept. 3, 1999) (on file with author).

80.   *See* Richard H. Sander, *Experimenting with Class-Based Affirmative Action*, 47 J. LEGAL EDUC. 472, 493 tbl.8 (1997) (reporting that the standard deviation for all law school applicants nationwide in the 1995-96 cycle was 8.0 points). Using standard deviation units permits comparisons across different scoring scales. For a definition, see LAW SCH. ADMISSION COUNCIL, THE LAW SCHOOL ADMISSION TEST: SOURCES, CONTENTS, USES 9 n.1 (1991).

> Standard deviation is a measure of the variation of a group of scores; that is, the larger the standard deviation, the more the individual scores differ from each other. It is calculated by squaring the difference between each score and the mean (i.e., squaring the "deviation"), adding together each of these squared deviations, and taking the square root of the sum. *One rough rule of thumb for standard deviations is that, in a normal population, about two thirds*

Since there were considerably more Asian Pacific Americans in the present investigation than in Gannon's previous research, this permitted separate reporting of the LSAT performance by Asian Pacific American subgroups. Too often, Asian Pacific Americans are either excluded from the affirmative action debate or "mascotted"[81] by affirmative action opponents in order to blunt the criticism that terminating affirmative action solidifies a system of racial privilege for Whites.[82] For example, professors Daniel Farber and Suzanna Sherry contend that the efforts of some feminists and critical race theorists to deconstruct merit are implicitly "anti-Asian."[83] Conservative historian Stephan Thernstrom argues that Asian Pacific American law school applicants have the most to gain by a ban on affirmative action in California.[84] Still other scholars assume that Asian

---

of the scores will be in the range of one standard deviation on either side of the mean while 95 percent of the scores will fall within the range of two standard deviations on either side of the mean.

Id. (italics added).

81. Professor Sumi Cho originated the term "racial mascot" to describe the political Right's positioning of Asian Pacific Americans within a larger agenda of obfuscating White hegemony. *See* Sumi K. Cho, *Redeeming Whiteness in the Shadow of Internment: Earl Warren,* Brown, *and a Theory of Racial Redemption,* 40 B.C. L. REV. 73, 169 (1998) ("The adoption of a racial group, or even an individual of color by a white political figure or constituency—a practice I refer to as mascotting—is necessary to deflect charges of racism and preserve the redeemed status of whiteness.").

82. For analysis of affirmative action, Asian Pacific Americans, and the "model minority" discourse, see generally Gabriel J. Chin et al., *Beyond Self-Interest: Asian Pacific Americans Toward a Community of Justice, A Policy Analysis of Affirmative Action,* 4 UCLA ASIAN PAC. AM. L.J. 129 (1996) (arguing in support of affirmative action and warning about the dangers of the model minority myth); Gautam Dutta, *Tokenism, 209, and the Politicization of Asian Americans,* 5 UCLA ASIAN PAC. AM. L.J. 45 (1998) (analyzing Asian American student and community activism in opposition to Proposition 209); Jerry Kang, *Negative Action against Asian Americans: The Internal Instability of Dworkin's Defense of Affirmative Action,* 31 HARV. C.R.-C.L. L. REV. 1 (1996) (arguing that Ronald Dworkin's justification for affirmative action can inadvertently permit discrimination against Asian Americans); Kidder, *Situating Asian Pacific Americans, supra* note 9, at 29 (reporting that Proposition 209 has had a negligible impact on Asian Pacific American enrollments at UC law schools, and arguing that affirmative action is not the major barrier to entry for Asian Pacific American law school applicants); Mari Matsuda, *We Will Not Be Used,* 1 UCLA ASIAN AM. PAC. ISLANDS L.J. 79 (1993) (arguing that affirmative action opponents have used Asian American success as a weapon for regressive politics); Frank H. Wu, *Neither Black nor White: Asian Americans and Affirmative Action,* 15 B.C. THIRD WORLD L.J. 225 (1995) (arguing that a thorough analysis of Asian Americans' position in U.S. society strengthens the case for affirmative action); Michael Omi & Dana Takagi, *Situating Asian Americans in the Political Discourse on Affirmative Action,* REPRESENTATIONS, Summer 1996, at 155 (critiquing both the Right and the Left discourse on Asian Americans and affirmative action).

83. *See* DANIEL A. FARBER & SUZANNA SHERRY, BEYOND ALL REASON 52-71 (1997).

84. *See* Thernstrom & Thernstrom, *supra* note 65, at 1629 ("The cost of racial double standards in admissions is currently being paid by many Asian students. When preferences are eliminated, they derive the greatest benefit. Thus Asian-American enrollment at the UCLA School of Law jumped by 73% when race-neutral admissions went into effect."). *But see* Kidder, *Situating Asian Pacific Americans, supra* note 9, at 38-41, 43-45 (reviewing 1994-99 data from UC law schools and concluding that Asian Pacific American enrollment, unlike White enrollment, did not rise noticeably as a result of ending affirmative action, and that Thernstrom employed misleading figures to arrive at the opposite conclusion).

Pacific Americans tend to do better than Whites on the LSAT.[85] Table 2 shows minority-White gaps on the LSAT for Asian Pacific American subgroups. The Asian Pacific American groups most underrepresented in higher education, Filipinos and Southeast Asians, tend to score over five points behind on the LSAT compared to Whites with equivalent UGPAs in the same school.

TABLE 2:
AVERAGE LSAT SCORE DIFFERENCES BETWEEN ASIAN PACIFIC
AMERICAN & WHITE
STUDENTS WITH EQUIVALENT PERFORMANCE IN THE SAME
COLLEGE[86]

Group	No. Matched	LSAT Gap (White - Minority)
Chinese	414	2.3
East Indian/Pakistani	167	0.7
Japanese	63	2.8
Korean	252	2.4
Filipino	61	5.5
Vietnamese, Thai, Cambodian & Laotian[87]	77	5.3

B. *Refined Matching: Controlling for Choice of Major*

One problem with matching students by UGPA within the same institution is that there still may be considerable differences in grading prac-

---

85. *See* Tomiko Brown-Nagin, *A Critique of Instrumental Rationality: Judicial Reasoning about the "Cold Numbers" in* Hopwood v. Texas, 16 LAW & INEQ. 359, 367 n.51 (1998). Brown-Nagin criticizes the *Hopwood* court's reasoning regarding the LSAT and Texas Index (TI) scores and notes:

> Since Asian-Americans often score higher than Whites on standardized measures of achievement, it is quite significant that the court excludes Asians from its analysis of minority/White differences in applicants' TI scores.... However, without the exclusion of Asian-Americans' scores, the court's claims concerning the singular harm to Whites of affirmative action for African-Americans and Latinos would have been undermined. That is, if the TI scores of Asian-American applicants tended to be higher than those of Whites, the court's understanding of who "steals" law school seats from deserving Whites would have been destabilized.

Id.

86. There were a handful of Polynesian students included in the overall matching, but they were excluded from Table 3 due to extremely small sample size (n=5).

87. These groups of Southeast Asians are clustered together because that is the classification practice of the Boalt Hall Admissions Office, and it was not possible to disaggregate the data.

tices within institutions.[88] For example, if, as some scholars suggest, students of color are disproportionately concentrated in "easier" majors (that is, majors with relatively lax grading requirements), then the matching analysis conducted heretofore could be misleading. It might exaggerate racial and ethnic differences on the LSAT among "equally performing" students because of the inadequacy of the way "equally performing" was defined, or so the argument goes.[89] Some may find it surprising that possible racial and ethnic differences in choice of major is such a contentious issue.[90] For example, a repeated claim of affirmative action critic Lino Graglia is that African Americans disproportionately enroll in non-competitive ethnic studies programs.[91] Similarly, Stephan Thernstrom responded with skepticism to a presentation of the preliminary results of this study at the Texas Legislature.[92] Thernstrom queried, "The question you have to ask is, were the students taking essentially the same courses?"[93] This is an important question as it relates to this study even though racial and ethnic differences in choice of major are not sufficient to support the

---

88.   *See* Robert L. Linn & C. Nicholas Hastings, *A Meta Analysis of the Validity of Predictors of Performance in Law School*, 21 J. EDUC. MEASUREMENT 245, 250 (1984). Linn and Hastings assert:

> First, there may be greater diversity in grading standards among colleges than among high schools. Second, law school applicants have a variety of undergraduate majors and the predictive meaning of the UGPA may vary across majors. Finally, the predictive value of the UGPA may be limited by so-called grade inflation.

Id. The issue of grade inflation is addressed in more depth in Part IV.D.1.

89.   *See* comments of Professor David Kaye *supra* Part I.C.

90.   For example, in 1998, UC Regent Ward Connerly proposed a review of ethnic studies departments and threatened to reduce or eliminate their funding. *See* John H. Bunzel, *Are Ethnic Studies Separate or Equal?*, L.A. TIMES, Nov. 8, 1998, at M2; Robin van der Vegt, *Ethnic Studies Create Debate*, NEWSDAY, Aug. 25, 1998, at A8; Joan Walsh, *Voices: Cal's Academic Question*, ASIAN WEEK, Sept. 23, 1998, at 5.

91.   *See, e.g.*, Lino A. Graglia, *Is Affirmative Action on the Way Out? Should It Be?*, COMMENTARY, Mar. 1998, at 31, 32 ("Inability to compete at the game being played necessarily leads to demands that the game be changed, and thus are born demands for black and ethnic studies and 'multiculturalism.' Such courses function to convince blacks that their academic difficulties are a result not of a lack of preparation but of white racism . . ."); Lino A. Graglia, *Racial Preferences in Admission to Institutions of Higher Education, in* THE IMPERILED ACADEMY 127, 135 (Howard Dickman ed., 1993).

> When the specially admitted students discover, as most soon must, that they cannot compete with their classmates, no matter how hard they try . . . they will insist, as self-respect requires, that the game be changed. Thus are born demands for black studies and multiculturalism, which perform the twin functions of reducing the need for ordinary academic work and providing support for the view that the academic difficulties of the black students are the result, not of substantially lower qualifications, but of racial antipathy—that the source of the problems is not black but white shortcomings.

Id.

92.   This presentation by David M. White, Director of Testing for the Public, included basic matching for Boalt's top five feeder institutions. Press Release, Testing for the Public, Top Minority Graduates Hurt by LSAT (Sept. 24, 1998) (on file with author).

93.   Brady R. Dewar & Vasant M. Kamath, *Report Shows LSAT Score Gap*, HARV. CRIMSON, Oct. 2, 1998, at 1.

contention of Graglia and others that the grades of students of color are "inflated."[94]

At first blush, the national evidence suggests that there is little reason to believe that race-related differences in choice of major have a substantial influence on UGPA (and LSAT score) differences. In a LSAC-sponsored study of the 1991 national law school applicant pool, Wightman found that "ethnic group membership is independent of undergraduate major category. Thus, there is no support for the hypothesis that members of certain ethnic groups tend to be clustered in specific undergraduate major areas."[95] Similarly, Bowen and Bok report that Black and White students in the College and Beyond Database generally chose the same array of traditional majors,[96] and that less than three percent of African Americans in highly selective schools major in African American studies, American Studies, or area studies of any kind.[97]

Nevertheless, with respect to assessing law school applicants, Wightman's, as well as Bowen and Bok's, general findings regarding field of study are not necessarily dispositive. For example, Wightman's findings are limited because she followed the usual practice of aggregating majors into several broad categories like arts and humanities, natural science, social science, health professions, and so forth.[98] In addition, a closer look at

---

94.   *Cf.* Sarah E. Turner & William G. Bowen, *Choice of Major: The Changing (Unchanging) Gender Gap*, 52 INDUS. & LAB. REL. REV. 289, 309 (1999).

> Since observed undergraduate performance measured by GPA is *conditional* on field choice and course selection, the observed relationship between undergraduate GPA and SAT scores will not provide an answer to the counterfactual question of how a student who did not major in a particular field would be expected to perform in that subject.

Id. A hypothetical example clarifies how atmospheric differences, rather than grade inflation, could produce the differences Graglia criticizes. Suppose that one compared students in two majors at the same college and it was possible to control for what is usually thought of as measuring prior educational attainment (high school GPA and SAT scores, for instance). And suppose that students of color majoring in ethnic studies received, on average, 3.20 UGPAs, and that equivalently prepared students of color majoring in history received 3.00 UGPAs. While critics like Graglia might argue that this is persuasive evidence of "grade inflation" and even a corruption of intellectual standards, the very same evidence could support a different environmental hypothesis. Suppose that there were something in the intellectual atmosphere of an ethnic studies program that, at least partially, removed the persistent undercurrent of racism faced by students of color in higher education. It would then follow that scholars using measures like the SAT or LSAT as metrics to assess "grade inflation" would be inculcating a race-related atmospheric bias into their supposedly neutral benchmark for comparing grades.

95.   Linda F. Wightman, *The Threat to Diversity in Legal Education: An Empirical Analysis of the Consequences of Abandoning Race as a Factor in Law School Admission Decisions*, 72 N.Y.U. L. REV. 1, 47 (1997) [hereinafter, Wightman, *Threat to Diversity*].

96.   *See* BOWEN & BOK, *supra* note 63, at 259 n.4. The majors in Bowen and Bok's database are individually listed in Appendix B, *id.* at 354. The dozens of majors fall into the following general categories: humanities, social science, natural science, engineering and other professional. *Id.*

97.   *Id.* at 71-72.

98.   *See* LINDA F. WIGHTMAN, LAW SCH. ADMISSION COUNCIL, WOMEN IN LEGAL EDUCATION 17 (1996); Wightman, *Threat to Diversity, supra* note 95, at 47. For example, in a study of gender and choice of major, Turner and Bowen decided to disaggregate social science majors because there were,

Bowen and Bok's data indicates that there are noticeable Black-White differences among top law school feeder majors. On the one hand, Black students in Bowen and Bok's College and Beyond database were more likely than Whites to major in political science, sociology and psychology.[99] On the other hand, White students were more likely than Blacks to major in history and English.[100] Since among all Boalt applicants English, history, political science, psychology and sociology were all common feeder majors, these findings are not inconsequential for purposes of the present matching analysis. It is important to control for specific majors to study how much this source of within-school variation influences racial or ethnic LSAT score differences among applicants with equivalent UGPAs in the same college.

In light of the aforementioned considerations, a second round of "refined" matching was conducted within each of the most common majors, including political science or government, history, economics, English, philosophy, psychology, and sociology. This control is in addition to those already implemented for college attended and graduation date. Table 3 shows that the ethnic gaps on the LSAT among students who perform similarly in college persist when comparing students in the same major. Comparing Table 1 and Table 3 demonstrates that for African Americans and Latinos, controlling for choice of major had a negligible impact on the size the LSAT gap among applicants with equivalent UGPAs. For African Americans the gap went from 9.2 points to 9.1 points, and for Latino applicants the gap increased from 6.8 points to 7.0 points. For Asian Pacific Americans, the gap actually increased from 2.5 points to 3.6 points when matching was conducted within the same major.[101]

---

for example, "pronounced gender-related differences between economics and psychology, and between each of these fields and the other social sciences." Turner & Bowen, *supra* note 87, at 292.

   99.   *See* BOWEN & BOK, *supra* note 63, at 71 fig.3.9.

   100.   *See id.*

   101.   The reasons for this increase are unclear, and are beyond the scope of the present investigation. It could mean that there were proportionately more Asian Pacific American law school applicants from majors with stiffer grading standards.

TABLE 3:
AVERAGE LSAT SCORE DIFFERENCES BY RACE/ETHNICITY
STUDENTS WITH EQUIVALENT PERFORMANCE[102] IN THE SAME
COLLEGE & SAME MAJOR

Racial/Ethnic Group	African American	Chicano/ Latino	Native American[103]	Asian Pacific American
No. Matched	120	249	—	540
LSAT Gap (White - Minority)	9.1	7.0	—	3.6

IV
DISCUSSION

While the results of this study document that the LSAT favors Whites among equally achieving college students, many leading scholars of affirmative action do not vigorously investigate whether standardized tests like the LSAT or SAT are biased against students of color. For example, in *The Shape of the River*, William Bowen and Derek Bok state: "We have no desire to enter the dense thickets of controversy surrounding the use of standardized tests."[104] Underlying this lack of interest in the test bias issue is a consensus among otherwise sharply divided scholars that racial/ethnic differences in LSAT scores reflect real underlying differences in academic or cognitive skills. This generalization holds for scholars like Charles Murray[105] and Arthur Jensen,[106] who assert that racial differences in cognitive ability are largely rooted in heredity, as well as conservative opponents of affirmative action like Stephan and Abigail Thernstrom[107] and Barbara

---

102. Consistent with the earlier matching, a +/-.10 UGPA bandwidth was employed.

103. Native Americans were excluded from the refined matching analysis because the sample was so small (n=33).

104. BOWEN & BOK, *supra* note 63, at 262 n.10.

105. RICHARD J. HERRNSTEIN & CHARLES MURRAY, THE BELL CURVE: INTELLIGENCE AND CLASS STRUCTURE IN AMERICAN LIFE (1994). For collections of differing views on Herrnstein and Murray's work, see THE BELL CURVE DEBATE (Russel Jacoby & Naomi Glauberman eds., 1995); THE BELL CURVE WARS (Steven Fraser ed., 1995). For a more focused critique, see INEQUALITY BY DESIGN: CRACKING THE BELL CURVE MYTH (Claude S. Fischer et al. eds., 1996).

106. ARTHUR R. JENSEN, BIAS IN MENTAL TESTING (1980); Arthur R. Jensen, *How Much Can We Boost I.Q. and Scholastic Achievement?*, 39 HARV. EDUC. REV. 1 (1969). For debate over Jensen's claims about race and IQ, see ARTHUR JENSEN: CONSENSUS AND CONTROVERSY (Sohan Modgil & Celia Modgil eds., 1987). For a critique of Jensen's claims about standardized tests, see Lorrie A. Shepard, *The Case for Bias in Tests of Achievement and Scholastic Aptitude, in* ARTHUR JENSEN: CONSENSUS AND CONTROVERSY, *supra*, at 177; Anne L. Hafner & David M. White, *Bias in Mental Research*, 51 HARV. EDUC. REV. 577 (1981) (book review).

107. THERNSTROM & THERNSTROM, *supra* note 50, at 402-03.

Lerner,[108] who believe that test score gaps accurately reflect differences in acquired academic skills rather than genetically-linked differences. It holds equally true for elite establishment proponents of affirmative action like Bowen and Bok.[109] Perhaps Philip Shelton, President of LSAC, best summarizes this conventional wisdom about the LSAT. Shelton claims, "[G]roup differences in performance on the LSAT almost certainly reflect differences in opportunities to acquire those skills. That is, these differences are real—they are not an artifact of the test itself."[110]

Here in Part IV, I argue that the LSAT is culturally biased because it artificially exaggerates educational differences between Whites and students of color. I attempt to explain how the LSAT can be biased even

---

108. Barbara Lerner, *The War on Testing: Detroit Edison in Perspective*, 33 PERSONNEL PSYCHOL. 11, 11 (1980).

> I believe that the answer is this: tests are under attack today because they tell us truths about ourselves and our society; partial truths, to be sure, but truths nonetheless and, in recent years, many of these truths have been unpleasant and unflattering. Seen in this perspective, the attack on tests is, to a very considerable and very frightening degree, an attack on truth itself by those who deal with unpleasant and unflattering truths by denying them and by attacking and trying to destroy the evidence for them.

Id. Lerner is presently an expert witness on standardized tests for the Center for Individual Rights in *Gratz v. Bollinger*, the challenge to the University of Michigan's undergraduate affirmative action program.

109. *See* BOWEN & BOK, *supra* note 63, at 262 n.10. Bowen and Bok conclude that the SAT is not biased against students of color:

> These findings [about minority underperformance] have an important bearing on the argument made by some proponents of race-sensitive admissions that differences in test scores should be disregarded because standardized tests, such as the SAT, are culturally biased against members of minority groups. To buttress this claim, opponents of these tests often call attention to particular questions that seem to call for familiarity with words or phrases that have little or no currency in poor urban neighborhoods or minority communities.... What is clear is that the evidence cited here shows that, far from being biased *against* minority students, standardized admissions tests consistently predict higher levels of academic performance than most blacks actually achieve.

Id. Arguments regarding overprediction are critiqued *infra* Part IV.C.1.

110. Philip D. Shelton, *"Top Ten" Misconceptions about the LSAT*, LAW SERVICES REP., Jan./Feb. 1999, at 4. This view echoes LSAC's position twenty years earlier in the *Bakke* case. While quick to refute claims about the hereditary origins of score differences, LSAC nonetheless treats gaps on the LSAT as reflective of genuine differences in skills. *See* Brief of Amicus Curiae Law School Admission Council, Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265 (1978), *reprinted in* ALLAN BAKKE VERSUS REGENTS OF THE UNIVERSITY OF CALIFORNIA: THE SUPREME COURT OF THE UNITED STATES VOLUME IV 143, 164 (Alfred A. Slocum ed., 1978).

> It would be contrary to basic premises of equality to suppose that a paper and pencil test of educational attainment could determine skin color among students who have been equally educated. Like college grades, test scores penalize blacks not because the tests measure innate intelligence or mental capacity, but rather because they measure abilities which are taught, acquired, and developed in formal education. A different, inferior education naturally tends to produce different, inferior scores.

Id.

In *Bakke*, a fractured Supreme Court struck down the affirmative action program at the UC Davis Medical School though it upheld the use of race as a plus factor in admission decisions. Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265, 320 (1978) (opinion of Powell, J.). For analysis of the viability of *Bakke* in light of the Court's more recent decisions, see Akhil Reed Amar & Neal Kumar Katyal, *Bakke's Fate*, 43 UCLA L. REV. 1745 (1996).

though my conclusion contradicts the consensus view of psychometricians and education scholars. 1 first show how the results of this study are highly consistent with Joseph Gannon's similar research two decades ago.

Second, I review the research on "stereotype threat" as a likely explanation of my findings. I reason that social stereotypes depress the standardized test performance of high-achieving students of color, particularly African Americans.

Third, I explain why I adopt an unorthodox matching procedure rather than the traditional approach of validating the LSAT as a predictor of first-year law school grades. I review LSAC research spanning several decades to reveal fundamental flaws in the conventional wisdom. On the one hand, LSAC researchers argue that the LSAT is not biased against minorities because the test predicts higher first-year law school grades than those students actually obtain. On the other hand, I demonstrate that these very same studies reveal consistently smaller White-minority differences in first-year law school grades compared to LSAT scores, a fact that is routinely obscured. I also argue that predictive validity evidence is inappropriately deployed by testing advocates to justify logically separate claims about selection fairness. As part of my critique of the conventional wisdom, I point out the narrowness of validating the LSAT against first-year law school grades, as opposed to second-and third-year grades or success in legal practice. I also explore evidence suggesting that the environment of legal education is contaminated by racial bias, which calls into question the neutrality of using law school grades to validate the LSAT. I next look at the possibility of biased content on the LSAT. I argue that the usual method to identify biased items, called differential item functioning, is tautological because its proponents assume that the test is fair overall in order to "prove" that individual questions, on balance, are fair.

Fourth, I evaluate the merits of three likely alternative explanations of the results. While college grade inflation ("compression") is a favorite jeremiad of educators, I find that the evidence overall fails to substantiate that grade compression undermines the reliability of UGPA as a matching criterion. I then analyze claims that UGPAs are biased in favor of students of color. I find that this hypothesis lacks empirical support and that there is more evidence substantiating the claim that institutional racism causes UGPAs to, if anything, be biased against minority students. Finally, I assess the objection that, in the absence of LSAT bias, it is still unsound to assume that racial and ethnic groups should perform similarly on the LSAT if matched on UGPA.

A. Comparison with Gannon's Earlier Matching Study

In 1981, Boston College statistician Joseph Gannon and the National Conference of Black Lawyers published a test bias study that departed

from the conventional predictive validity approach.[111] Rather, Gannon adopted a "post-dictive"[112] methodology, meaning he looked backward in time to see if he could confirm the hypothesis that Whites and students of color with equivalent academic performance over four years of college would likewise obtain similar scores on the LSAT.[113] This study is modeled after Gannon's research.

Gannon obtained data on the four colleges that provided the most applicants to each of a dozen ABA-accredited law schools, including six public and six private, that varied in both geography and prestige.[114] He matched White applicants with minority students who possessed UGPAs within +/-.10 from the same college.[115] On the old 200-800 scale, Gannon found that African American students still trailed their equally accomplished White counterparts on the LSAT by 110 points (approximately one standard deviation) after controlling for equivalent performance in the same college.[116] Furthermore, Latino students trailed by 97 points, Native American students by 78 points, and Asian Pacific American students by 36 points.[117]

Before concluding that there was some sort of bias on the LSAT or in the testing atmosphere, Gannon evaluated several alternative hypotheses. First, Gannon investigated whether differences in grading standards over time (that is, grade inflation) accounted for the different results. Second, he considered whether variations in the academic rigor of different fields of study may have explained the racial LSAT performance differences among students with similar UGPAs from the same institution.[118] Gannon controlled for graduation date and field of study for the portion of his database where this information was available.[119] As with the current study, after

---

111. Gannon, *supra* note 5. Typically researchers employ regression techniques to assess the strength of the correlation between the LSAT and first year law student grades. For more detail, see infra Part IV.C.

112. For another example of a "post-dictive" study, see Donald E. Powers, *Effects of Test Preparation on the Validity of a Graduate Admissions Test*, 9 APPLIED PSYCHOL. MEASUREMENT 179 (1985) (connecting hours of coaching for the GRE with UGPAs rather than graduate school grades).

113. Gannon, *supra* note 5, at 273.

114. *Id.* at 274-75. The total sample size for Gannon's database was 19,287, which made it possible to match 727 African Americans, 482 Latinos, 379 Asian Americans, and 48 Native Americans. *Id.* The data request letter sent by the National Conference of Black Lawyers to the law schools stipulated that the study would preserve institutional confidentiality, so the names of the law schools and undergraduate feeder institutions are unavailable. The present study included no such agreement.

115. *Id.* at 276.

116. *Id.* at 276-77. Prior to 1982 the LSAT was on a 200-800 scale, then it was on a 10-48 scale from 1982-1991, and it has been on a 120-180 scale since June 1991. LAW SCHOOL ADMISSION COUNCIL, *supra* note 80, at 2, 10.

117. Gannon, *supra* note 5, at 276-77.

118. *Id.* at 278.

119. *Id.* One of the twelve law schools in Gannon's study provided information on majors. *Id.* at 278.

this "refined" round of matching, Gannon found that adding controls for graduation date and choice of major had little effect on the LSAT gaps.[120] Gannon also tested whether the dramatic gaps on the LSAT were partly attributable to statistical skewing that resulted from minority applicants in his sample having lower UGPAs overall.[121] Gannon found that group differences in UGPA distributions do not create a sample-selection bias that exaggerates group differences in LSAT averages, despite rigorous efforts at matching.[122]

Gannon concluded there was ample support for the hypothesis that the LSAT or the testing milieu was biased against students of color. In other words, minority law school applicants faced a LSAT bias *in addition to* disadvantages in prior educational opportunities. Gannon argued:

> Clearly, the data reveal that law school applicants with essentially equivalent college grades are apt to receive widely discrepant LSAT scores depending upon their race or ethnicity. Whereas past discrimination and deprivation must always be acknowledged, it is important to realize that this study pertains to minority applicants who overcame innumerable social and economic barriers as evidenced by the criterion of academic parity in matching them with majority applicants in the same applicant pool. Consequently, academic equals are suddenly and systematically reduced to intellectual inferiors as a result of their poorer performance in the testing situation, or so the argument goes for those believing that the test differences are truly "real." Put rather bluntly, what it took four years to build up, it required roughly four hours to tear down.[123]

Until now, Gannon's study, which is twenty years old, has never been confirmed or challenged by replication. There are three additional reasons it was important to revisit the issues raised by Gannon's study. First, the national law school applicant pool underwent major changes. In 1978-79, 8565 students of color applied to ABA-accredited law schools,[124] compared to 20,300 in 1998-99.[125] The current study disproves the possibility that the

---

120.   *Id.* at 280.

121.   *Id.*

122.   *Id.* After conducting regression analysis, Gannon found that the correlation coefficients were in the direction consistent with this sample-selection bias hypothesis, but were insufficient in magnitude to have much of an impact on LSAT score differences. *Id.* Rather, Gannon found the LSAT discrepancies among matched applicants that substantially favored Whites "persist across the entire range of the grading curve." *Id.*

123.   *Id.* at 282.

124.   LAW SCH. ADMISSION COUNCIL, THE CHALLENGE OF MINORITY ENROLLMENT tbl. 16 (1981). These figures included Black, Chicano, and "unspecified minority." *Id.* At the time there were 60,280 White and "unidentified" applicants. *Id.*

125.   LAW SCH. ADMISSION COUNCIL, 1998-99 NATIONAL DECISION PROFILES (2000) [hereinafter DECISION PROFILES]. The 20,300 figure is a composite of Native American, Black, Asian Pacific American, Chicano, other Hispanic, and Puerto Rican applicants. *Id.* White applicants, by comparison, dropped to 47,787. *Id.*

LSAT gaps reported by Gannon were at least partly attributable to a temporary "trailblazer effect." Second, the content and format of the LSAT was substantially revised in 1982, 1989, and 1991,[126] raising questions about the general applicability of Gannon's findings to today's admissions environment. Finally and most importantly, test bias can be a potential justification for continuing affirmative action.[127]

B.   Stereotype Threat Research: A New Perspective on an Enduring Debate

The scholarly discourse on test bias acknowledges that historically marginalized groups may face added pressure and anxiety that disproportionately depresses their performance.[128] Yet, it is also true that the last five years have ushered in a new line of provocative research regarding the psychosocial atmosphere of standardized test taking. Whereas LSAC and other test producers generally assume that giving a standardized test to students under identical testing conditions ensures that the test is a neutral, objective measure of abilities for different groups,[129] psychologists Claude Steele, Joshua Aronson, and their colleagues demonstrate that the milieu of standardized test taking can be experienced differently depending on one's race or ethnicity.[130] In other words, merely making the content of the test the

---

126.   *See* LAW SCHOOL ADMISSION COUNCIL, *supra* note 80 at 3-5 (reviewing historical changes in the LSAT).

127.   *See* Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265, 306 n.43 (1978).

> Racial classifications in admissions conceivably could serve a fifth purpose, one which petitioner does not articulate: fair appraisal of each individual's academic promise in the light of some cultural bias in grading or testing procedures. To the extent that race and ethnic background were considered only to the extent of curing established inaccuracies in predicting academic performance, it might be argued that there is no 'preference' at all.

Id.; Grutter v. Bollinger, 188 F.3d 394, 401 (6th Cir. 1999) (permitting intervention by student of color organizations after reasoning: "We find persuasive their argument that the University is unlikely to present evidence of past discrimination by the University itself or of the disparate impact of some current admissions criteria, and that these may be important and relevant factors in determining the legality of a race-conscious admissions policy."); *cf.* DeFunis v. Odegaard, 416 U.S. 312, 335 (1974) (Douglas, J., dissenting) ("My reaction is that the presence of an LSAT is sufficient warrant for a school to put racial minorities into a separate class in order to better probe their capacities and potentials.").

128.   Ray Hembree, *Correlates, Causes, Effects and Treatment of Test Anxiety*, 58 REV. EDUC. 47, 61 (1988) (summarizing research literature on gender and test anxiety).

129.   *See* LAW SCH. ADMISSION COUNCIL, *supra* note 80, at 26 ("Because the LSAT is administered to all applicants under standard conditions and each test form requires the same or equivalent task of everyone, LSAT scores provide a standard measure of abilities."); LAW SCH. ADMISSION COUNCIL, LAW SCHOOL ADMISSION REFERENCE MANUAL, 3.4 (1995-96) ("The LSAT permits the direct comparison of the abilities of persons from diverse educational backgrounds. . . . The primary advantage is that [LSAT scores] provide a standard measure and are administered to all applicants under standard conditions.").

130.   *See* Claude M. Steele, *Thin Ice: "Stereotype Threat" and Black College Students*, ATLANTIC MONTHLY, Aug. 1999, at 44, 47 [hereinafter Steele, *Thin Ice*]. Steele writes:

> In matters of race we often assume that when a situation is objectively the same for different groups, it is *experienced* in the same way by each group. . . . But for black students, difficulty

same for everyone does not guarantee that taking the test will be the same regardless of race or ethnicity.[131]

Steele and Aronson use the term "stereotype threat" to describe how the threat of being viewed through the lens of a negative stereotype about one's group can interfere with performance on standardized tests by members of groups who tend to do less well on such tests. This phenomenon is not necessarily linked to the cultural attributes of any particular marginalized group. Rather, stereotype threat can affect any group where there exists a widely recognized negative stereotype about that group's performance in a certain domain.[132]

Steele and Aronson administered a test of challenging Graduate Record Examination (GRE) questions to Black and White college students who were matched on SAT scores. When it was suggested that the test measures ability, and this suggestion is, after all, implicitly made whenever students take a high-stakes test like the LSAT or SAT, Black performance was lower than that of comparable Whites.[133] Steele and Aronson reasoned that if stereotype threat per se depresses performance, then altering testing conditions so as to remove that stereotype threat should cause performance to improve.[134] Thus, they administered the very same GRE test to another cohort of Black and White college students matched on SAT scores, only this time they presented the test as a laboratory problem-solving task that was not indicative of ability.[135] By altering the social meaning of the test, and keeping stereotype threat at bay, Black performance suddenly equaled that of Whites.[136] Whites performed no worse (actually they did insignificantly better) when the test was labeled as indicative of ability because, unlike African Americans, Whites are not under the stereotype threat spot-

---

> with the test makes the negative stereotype relevant as an interpretation of their performance, and of them. They know that they are especially likely to be seen as having limited ability. Groups not stereotyped in this way don't experience this extra intimidation. And it is a serious intimidation, implying as it does that they may not belong in walks of life where the tested abilities are important—walks of life in which they are heavily invested. Like many pressures, it may not be experienced in a fully conscious way, but it may impair their best thinking.

Id.

131. Expert Report of David M. White, *supra* note 45, at 17 ("Nevertheless, the different ways students take tests can affect their scores without reflecting their underlying abilities to read or reason. Simply giving the same test to all applicants cannot ensure that all applicants will take the same test.").

132. Claude M. Steele, *A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance*, AM. PSYCHOL., June 1997, at 613, 617 [hereinafter Steele, *A Threat in the Air*].

133. Claude M. Steele & Joshua Aronson, *Stereotype Threat and the Intellectual Test Performance of African Americans*, 69 J. PERSONALITY & SOC. PSYCHOL. 797, 799-801 (1995). This study involved Stanford undergraduates.

134. Expert Report of Claude M. Steele, Grutter v. Bollinger, 16 F. Supp. 2d 797 (E.D. Mich. 1998) (No. 97-75928). Steele's report is reprinted in 5 MICH. J. RACE & L. 439 (1999). All subsequent references are to the reprinted version.

135. Steele & Aronson, *supra* note 133, at 799.

136. *Id.* at 800 fig.1.

light in the domain of standardized test taking.[137] Announcing the test as diagnostic of ability impaired both the accuracy and the rate at which African Americans completed the test.[138]

Similar results occur with respect to the standardized test performance of other populations vulnerable to a societal stereotype about the performance of their group, including Latinos on a difficult English test,[139] women on an advanced mathematics test,[140] and low-income French students on a difficult language test.[141] Indeed, when White males with high SAT math scores were given a challenging mathematics test of GRE questions and were told that the study's purpose was to understand why Asians perform better than Whites in mathematics,[142] White male performance dropped significantly, while it remained equal in the non-diagnostic cohort.[143]

Steele and Aronson were able to show further that race-related anxiety is specifically evoked when stereotype threat undermines the standardized test performance of African Americans. Not only did Steele and Aronson demonstrate that merely presenting a GRE verbal test as diagnostic of abilities increased anxiety levels[144] and blood pressure,[145] they found that African Americans in the diagnostic condition were more cognizant of race-related issues and displayed higher levels of self-doubt, based on

---

137.   *See id.*

138.   *Id.* at 802.

139.   *See* Joshua Aronson et al., The Effect of Stereotype Threat on the Standardized Test Performance of College Students, Paper Presented at the Annual Convention of the American Psychological Society 10-12 (May 1998) (unpublished manuscript, on file with author) (reporting the results of stereotype threat research at the University of Texas-Austin using GRE questions).

140.   *See* Steven J. Spencer et al., *Stereotype Threat and Women's Math Performance*, 35 J. EXPERIMENTAL SOC. PSYCHOL. 4 (1999) (finding stereotype threat depressed women's math performance among groups of University of Michigan students, while removing stereotype threat equalized performance).

141.   *See* Jean-Claude Croizet & Theresa Claire, *Extending the Concept of Stereotype Threat to Social Class: The Intellectual Under-Performance of Students from Low Socioeconomic Backgrounds*, 24 PERSONALITY & SOC. PSYCHOL. BULL. 588 (1998).

142.   *Id.* at 37.

143.   *See* Joshua Aronson et al., *When White Men Can't Do Math: Necessary and Sufficient Factors in Stereotype Threat*, 35 J. EXPERIMENTAL SOC. PSYCHOL. 29 (1999) (describing research at Stanford and the University of Texas-Austin). Like earlier stereotype threat research, these students were matched on SAT scores. Also note that this finding is not inconsistent with the present finding that Asian Pacific Americans had lower LSAT scores than Whites. There is no necessary inconsistency because the stereotype about "Asian test takers" appears to be more pronounced with respect to mathematics, which is not a part of the LSAT. In addition, this stereotype may not have the same level of across-the-board salience as the stereotype faced by African American students.

144.   Steele & Aronson, *supra* note 133, at 801-02 (reporting statistically significant differences for the Black diagnostic and non-diagnostic cohorts when the Speilberger State Anxiety Inventory was administered immediately after the test).

145.   Steele, *Thin Ice, supra* note 130, at 51 (describing research by James Blascovich of UC Santa Barbara, who found that blood pressure of Black students performing a challenging cognitive task under stereotype threat was higher than that under the non-diagnostic condition, and higher than that of Whites under either condition).

word-fragment completion tests given after the test instructions.[146] They also were less likely to report participating in activities and interests that are commonly associated with African American culture. That is, the Black students in the diagnostic condition were actively avoiding conformity with stereotypic images, even while their fear of conforming to a negative group stereotype lowered their test performance. For White students, announcing the test as either diagnostic or non-diagnostic of ability had no significant effect on blood pressure, self-doubt, or the cognitive salience of race-related issues.[147]

There are two other important issues relating to stereotype threat research. The first point is that while Steele and Aronson's methods for researching stereotype threat are quite sophisticated in their detection and measurement of the psychosocial dynamics of test taking, they are not able to capture the cumulative effects of this bias over time as students navigate through one standardized test after another. While stereotype threat provides an important snapshot of how bias is created in the milieu of standardized testing, it would be fruitful (though exceedingly difficult) to attempt to longitudinally track the effects of stereotype threat over time.[148]

The second point is that stereotype threat is not the product of a lack of motivation or effort. Quite the opposite, Steele argnes:

> Stereotype threat is especially frustrating because, at each level of schooling, it affects the vanguard of these groups, those with the skills and self-confidence to have identified with the domain. Ironically, their susceptibility to this threat derives not from internal doubts about their ability (e.g., their internalization of the stereotype) but from their identification with the domain and the resulting concern they have about being stereotyped in it.[149]

---

146. Steele & Aronson, *supra* note 133, at 803-04.

147. *Id.* at 805 fig.3; Steele, *Thin Ice, supra* note 130, at 51.

148. This Comment is not intended as a criticism of Steele, Aronson, and their colleagues. Rather, it is a reminder to policy makers (and an invitation to researchers) that, like differences in wealth, the effects of test bias can accumulate over time, creating increasing barriers to opportunity as one proceeds along the higher education path. Elsewhere 1 have analyzed the cumulative effects of gender bias on standardized testing among the same cohort between the five-year span of taking the SAT, attending college, and then taking the GRE. Kidder, *Portia Denied, supra* note 47, at 26-28. However, it is difficult to conclude with certainty that the cause of such a bias is stereotype threat per se, or some other source.

149. Steele, *A Threat in the Air, supra* note 132, at 614; *see also Affirmative Action and Standardized Testing,* 4 TEX. HISP. J.L. & POL'Y 85, 93 (1998). At a panel discussion, Professor Joshua Aronson commented:

> The more and more I've done this, the more I've started to lose my faith in the standardized testing paradigm, and here's why. . . . In just about every one of these studies, we find that people who care more about the domain, actually performed worse in a stereotype threat situation. When we asked them how much do you care about math, for example, it's the women who most care about math, who most underperform in these situations where they feel like they're being compared to males. Think about that for a second, because if you really want to look at who the good students are in college, probably the best predictor is how much they care about succeeding, how much do they love the subject that they're involved in?

Thus, the stereotype threat research is especially helpful in illuminating the present findings, which involved LSAT gaps among equally performing elite college students applying to a highly competitive law school.

If stereotype threat vanquishes the "vanguard," the most accomplished underrepresented minority students, this finding suggests that high-achieving underrepresented minorities pay what Steele calls a "pioneer tax"[150] when taking standardized tests like the LSAT. Similarly, social psychologist Sherman James describes the psychological price Blacks pay for exerting enormous effort in the face of obstacles like stereotype threat as "John Henryism."[151] Thus, admissions policies that attempt to replace race-sensitive affirmative action with class-based affirmative action[152] will unfairly overlook the disadvantages middle class African Americans[153] face in competing for spots in selective higher education institutions where standardized tests figure prominently. The present LSAT study makes clear how the "John Henryism" of building up a record of equal performance during four plus years of college can be dismantled in the time it takes to complete a half-day standardized test.[154]

C. Failures of the Conventional Wisdom about Test Bias

One of the first questions many readers, and especially those steeped in the testing literature, are likely to ask is this: Why did 1 employ a "post-dictive" matching approach to the study of cultural bias on the LSAT, rather than examine whether the LSAT is equally effective at predicting law school grades for various racial and ethnic groups? After presenting background information on the predictive validity approach, I show how

---

Id.

150.   Expert Report of Claude M. Steele, *supra* note 134, at 447 (arguing that minority students invested in high achievement pay a "pioneer tax" of worry and vigilance because of concern about the societal treatment of their group).

151.   Steele, *Thin Ice, supra* note 130, at 51. John Henry is a Black figure in American folklore, famed for his ultimately fatal heroism in competing against a steam drill to finish digging a railroad tunnel. For background material, see Brett Williams, JOHN HENRY: A BIO-BIBLIOGRAPHY (1983).

152.   For arguments regarding socioeconomic-based affirmative action, see Richard D. Kahlenberg, *The Remedy: Class-Based Affirmative Action*, 84 CALIF. L. REV. 1037 (1996); Deborah C. Malamud, *Assessing Class-Based Affirmative Action*, 47 J. LEGAL EDUC. 452 (1997); Deborah C. Malamud, *Class-Based Affirmative Action: Lessons and Caveats*, 74 TEX. L. REV. 1847 (1996); Sander, *supra* note 80, at 472; Deborah C. Malamud, *A Response to Professor Sander*, 47 J. LEGAL EDUC. 504 (1997); Richard H. Sander, *Comment in Reply*, 47 J. LEGAL EDUC. 512 (1997); Richard D. Kahlenberg, *In Search of Fairness: A Better Way*, WASH. MONTHLY, June 1998, at 26.

153.   For an analysis of the continuing barriers to opportunity faced by middle class Blacks, see Kidder, *Rise of the Testocracy, supra* note 47, at 182-86 (discussing how Black-White differences in net worth are far greater than income differences, and how definitions of socioeconomic disadvantage that are based on income thus siguificantly underestimate the true magnitude of Blacks' disadvantage vis-à-vis Whites); Deborah C. Malamud, *Affirmative Action, Diversity, and the Black Middle Class*, 68 U. COLO. L. REV. 939 (1997) (detailing the race-based economic inequality encountered by middle class African Americans).

154.   *See* text accompanying *supra* note 123.

this conventional method creates a misleading picture of minority under-performance, even though there is a smaller gap between Whites and students of color in law school performance than in LSAT performance. 1 attempt to demonstrate how this predictive validity evidence is then improperly used to buttress claims that the LSAT is a racially unbiased selection device. I also investigate problems with first-year law school grades, which is the criterion against which the LSAT is typically validated. Finally, I identify problems with another common procedure called differential item functioming that is used to ensure cultural fairness on the LSAT.

The predictive validity (establishment) approach uses linear regression statistical techniques to calculate correlation coefficients between LSAT scores and first-year law school grade-point averages (FYAs).[155] Correlation coefficients provide a measure of the strength of the association between two variables, and can range from -1.0 to +1.0, with +1.0 indicating a perfect positive association, 0.0 indicating no relationship, and -1.0 indicating a perfect negative association.[156]

LSAC's recent predictive validity studies, which include almost all ABA-accredited law schools, suggest that the median correlation between the LSAT and FYAs is about +.40.[157] Since the square of the correlation coefficient provides a measure of the variance in the criterion (FYA) explained by the predictor (LSAT),[158] the LSAT accounts for about sixteen percent of the variance in first-year law school grades among those admitted to ABA-accredited law schools. Many scholars argue that the LSAT-FYA correlation is artificially depressed because of restriction of range;[159] that is, admitted law students have a narrower range of LSAT scores than a random population, or because of measurement error in the criterion variable of FYAs.[160] Thus, LSAC and other test developers

---

155. For a discussion of correlation coefficients and regression analysis in the context of the LSAT, see Lisa A. Anthony et al., Law Sch. Admission Council, Predictive Validity of the LSAT: A National Summary of the 1995-1996 Correlation Studies 6 tbl.2 (1999); Linda F. Wightman, Law Sch. Admission Council, Predictive Validity of the LSAT: A National Summary of the 1990-1992 Correlation Studies (1993); Kaye, *supra* note 54, at 447-54; Robert L. Linn, *Test Bias and the Prediction of Grades in Law School*, 27 J. Legal Educ. 293, 297-304 (1975); Linn & Hastings, *supra* note 88, at 507; W. B. Schrader, *Summary of Law School Validity Studies, 1948-1975, in* 3 Reports of LSAC Sponsored Research: 1975-1977, 519 (Law Sch. Admission Council ed., 1977); Michael Selmi, *Testing for Equality: Merit, Efficiency, and The Affirmative Action Debate*, 42 UCLA L. Rev. 1251, 1262-64 (1995); Sturm & Guinier, *supra* note 78, at 970-72; Wightman, *Threat to Diversity*, *supra* note 95, at 31-34.

156. Wightman, *Threat to Diversity*, *supra* note 95, at 31; Selmi, *supra* note 155, at 1263.

157. *See supra* note 155 and accompanying text for a summary of this literature.

158. Selmi, *supra* note 155, at 1263. For a detailed description of the concept of variance explained and regression, see Kaye, *supra* note 54, at 448-450.

159. *See* Anthony et al., *supra* note 155, at 10; Shelton, *supra* note 110, at 4; Wightman, *Threat to Diversity*, *supra* note 95, at 33.

160. Selmi, *supra* note 155, at 1268-69 (discussing and then criticizing the possibility that unreliability in grading diminishes predictive validity).

sometimes calculate "corrections" for range restriction,[161] a practice sometimes criticized for relying upon questionable assumptions.[162]

1. *Overprediction: Facts and Artifacts*

Questions about the overall predictive validity of a standardized test like the LSAT go hand in hand with concerns about a test's validity for racial and ethnic minority group applicants.[163] For instance, in a recent article defending standardized tests, Professor Dan Subotnik concludes, "Finally, and most important for our purposes, challenges to the validity of the LSAT based on minority performance are groundless."[164] Subotnik bases his strongly worded claim on LSAC studies.[165] LSAC (and psychometricians generally) traditionally address the fairness of tests like the LSAT by assessing whether the LSAT has what is known as "differential prediction" for students of color. That is, does the test predict lower FYAs than those students actually achieve?[166] LSAC's national studies consistently find that the LSAT and UGPA both tend to overpredict rather than underpredict the FYAs of African Americans and Latinos.[167] The finding

---

161.   *See, e.g.*, ANTHONY ET AL., *supra* note 155, at 10 (providing estimated correlation coefficients adjusted for restriction of range); Donald B. Rubin, *Using Empirical Bayes Techniques in the Law School Validity Studies, in* 4 REPORTS OF LSAC SPONSORED RESEARCH: 1978-1983, at 1, 17 (Law Sch. Admission Council ed., 1984) (providing estimated correlations between predictors and FYA for the entire applicant population rather than enrolled students); Linn & Hastings, *supra* note 88, at 536-41 (calculating various adjustments for range restriction).

162.   To an unknown degree, corrections for restriction of range will violate assumptions about linearity and homoscedasticity. *See* Kidder, *Rise of the Testocracy, supra* note 47, at 205-06; Selmi, *supra* note 155, at 1266-68; Sturm & Guinier, *supra* note 78, at 972-74.

163.   Wightman, *Threat to Diversity, supra* note 95, at 33 ("Questions about the overall validity of the LSAT often are raised in conjunction with concerns about its validity for racial or ethnic minority group applicants.").

164.   Dan Subotnik, *Goodbye to the SAT, LSAT? Hello to Equity by Lottery? Evaluating Lani Guinier's Plan for Ending Race Consciousness,* 43 How. L.J. 141, 153 (2000).

165.   *See id.* at 154 n.76.

166.   *See* AM. EDUC. RESEARCH ASS'N ET AL., STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING 79 (1999) ("Under one broadly accepted definition, no bias exists if the regression equations relating the test and the criterion are indistinguishable for the groups in question."). For a classic discussion of the related concepts of differential validity and differential prediction, see Robert L. Linn, *Single-Group Validity, Differential Validity, and Differential Prediction,* 63 J. APPLIED PSYCHOL. 507 (1978).

167.   *See* LINDA F. WIGHTMAN & DAVID G. MULLER, LAW SCH. ADMISSION COUNCIL, AN ANALYSIS OF DIFFERENTIAL VALIDITY AND DIFFERENTIAL PREDICTION FOR BLACK, MEXICAN AMERICAN, HISPANIC, AND WHITE LAW SCHOOL STUDENTS (1990); Donald E. Powers, *Comparing Predictions of Law School Performance for Black, Chicano, and White Law Students, in* 3 REPORTS OF LSAC SPONSORED RESEARCH: 1975-1977, *supra* note 155, at 721, 744; W.B. Schrader & Barbara Pitcher, *Predicting Law School Grades for Black American Law Students, in* 2 REPORTS OF LSAC SPONSORED RESEARCH: 1970-1974, 451, 453 (Law Sch. Admission Council ed., 1976); Wightman, *Threat to Diversity, supra* note 95, at 29 n.64, 34 (reviewing studies finding overprediction). Native Americans are consistently excluded from these studies due to insufficient samples, and Asian Americans have been historically excluded because of small samples and perhaps also a perception that since test score gaps between Whites and Asian Americans were much smaller, the adverse social

that the LSAT tends to predict higher FYAs than underrepresented minorities actually obtain is often cited by LSAC researchers to support the position that traditional law school admission criteria, like the LSAT and UGPA, do not unfairly burden minority applicants.[168] Moreover, affirmative action critics seize upon evidence of overprediction to claim that standardized tests are biased in favor of underrepresented minorities because students of color lag even further behind based on law school grades than on LSAT scores.[169]

Both of these related interpretations, however, are subject to criticism on several grounds. First, statistical artifacts associated with regression analysis can produce an appearance of minority overprediction that is not educationally meaningful, particularly when the overall correlation between the predictor and the criterion variable is low.[170] Second, and more importantly, critics scarcely acknowledge that even if the LSAT does genuinely overpredict minority FYAs within the regression framework, this overprediction still does not mean that students of color are doing

---

consequences that would attach to differential validity were not as substantial. For a discussion of law school admissions and Native American issues, see *supra* note 79.

168.   WIGHTMAN & MULLER, *supra* note 167, at 25 ("When a regression equation is developed using combined data from white and minority students, the equation tends to overpredict law school performance for minority students. There is nothing in these data to suggest that using the traditional predictors disadvantages minority law school applicants."); Powers, *supra* note 167, at 721, 748 (concluding from minority overprediction evidence that "[t]here was no evidence of unfairness, as viewed from the traditional regression model of bias"); Kristine S. Knaplund & Richard H. Sander, *The Art and Science of Academic Support*, 45 J. LEGAL EDUC. 157, 160 (1995). Knaplund and Sander suggest:

> It is often suggested that [LSAT disparities] exaggerate the actual gap in the ability of students to compete in law school because the LSAT is biased against minority students and understates their promise for achievement. But for statistical reasons (unrelated to culture) this argument is wrong, at least at UCLA: at a given LSAT level, the score overpredicts the performance of black students relative to that of whites.

Id. (footnote omitted).

169.   ROBERT KLITGAARD, CHOOSING ELITES 162-63 (1985) (using data from an LSAC study of differential validity to argue that 110 points should be taken away from Blacks' LSAT scores (on the 200-800 scale) if one wanted to correct for overprediction). Opponents of race-conscious affirmative action often repeat Klitgaard's argument. *See* Lino A. Graglia, Hopwood: *A Plea to End the "Affirmative Action" Fraud*, 2 TEX. F. ON C.L. & C.R. 105, 105 n.2 (1996); Stephan Thernstrom, *Diversity and Meritocracy in Legal Education: A Critical Evaluation of Linda F. Wightman's "The Threat to Diversity in Legal Education"*, 15 CONST. COMMENT. 11, 31 (1998). For a critique of Klitgaard, see David E. Van Zandt, *Merit at the Right Tail: Education and Elite Law School Admissions*, 64 TEX. L. REV. 1493 (1986) (book review); David M. White, *Two Views of Standardized Testing*, 55 HARV. EDUC. REV. 332 (1985) (book review).

170.   A full explanation of why statistical artifacts produce non-relevant overprediction is beyond the scope of the present inquiry. For an analysis of statistical artifacts' influence on differential validity, see UNIV. OF CAL., FINAL REPORT OF THE TASK FORCE ON GRADUATE AND PROFESSIONAL ADMISSIONS 43 (1977); Kidder, *Rise of the Testocracy*, *supra* note 47, at 182-89; Robert L. Linn & C. Nicholas Hastings, *Group Differentiated Prediction*, 8 APPLIED PSYCHOL. MEASUREMENT 165 *passim* (1984); White, *supra* note 169, at 335-36; Linda F. Wightman, *Are Other Things Essentially Equal? An Empirical Investigation of the Consequences of Including Race as a Factor in Law School Admission*, 28 Sw. U. L. REV. 1, 11 (1998).

relatively worse in law school than they perform on the LSAT. In this section, I will establish that the opposite is consistently true. A careful look at several of LSAC's differential validity studies indicates that African Americans, for example, differ less from White students in their FYAs than on the LSAT, notwithstanding the fact that those very same studies report that the LSAT (and UGPA) overpredict African Americans' FYAs.[171]

For example, Wightman and Muller's 1990 study of fifty-one law schools reveals that when Black-White gaps in FYAs and LSAT scores are converted to standard deviation units, which provides a uniform measure of dispersion around an average, African Americans are consistently closer based on FYAs than on LSAT scores.[172] Wightman's more recent analysis of "elite" law schools shows larger gaps on the LSAT than in actual FYAs for African Americans.[173] Powers's study for LSAC, which formed the basis of Klitgaard's conclusion that the LSAT was predictively biased in favor of Blacks, yields similar evidence of a smaller gap in FYAs than in LSAT scores,[174] as does Schrader and Pitcher's earlier research.[175] Though

---

171. *Infra* notes 172-175.

172. *See* WIGHTMAN & MULLER, *supra* note 167, at 9 tbl.6a. Wightman and Muller report on fifty-one schools in the 1986-88 period, where at least thirty African American students were enrolled, and provide the means and standard deviations for Whites' and Blacks' LSAT scores, UGPAs, and FYAs (converted to a uniform scale) at each law school. By comparing how much the standard deviation ranges overlap, one can assess the relative magnitude of the gaps on these different performance measures. *See* James C. Hathaway, *The Mythical Meritocracy of Law School Admissions*, 34 J. LEGAL EDUC. 86, 93 tbl.6 (1984) (reporting the degree of overlap between the upper limit of the minority standard deviation range and the lower limit of the White standard deviation range). Wightman and Muller's data reveal that at twenty-seven of the law schools the White and Black LSAT standard deviation ranges do not overlap, whereas the FYA ranges do not overlap at only five schools. *See* WIGHTMAN & MULLER, *supra* note 167, at 9 tbl.6a. Similarly, the upper limit of the Black standard deviation range reaches the White mean on the LSAT at only one law school, whereas this is the case at eleven law schools based on FYA. Overall, there is even less of a gap based on UGPA than on FYA (.95 standard deviations versus 1.25 standard deviations). *Id.* at 8. This may at least partly be caused by the greater traditionalism and subtle biases of legal education (especially in the first year) compared to undergraduate studies, where students have more autonomy to choose their course of study. For discussion of the literature on bias in higher education, see *infra* Part IV.C.3 and Part IV.D.2.

LSAC since has completed a more recent differential validity study. *See* Lisa C. Anthony et al., Law Sch. Admission Council, Analysis of Differential Prediction of Law School Performance by Ethnic Subgroups (forthcoming 2001). However, this LSAC report was not in print at the time this Comment was submitted for publication, and repeated efforts to procure an advance copy from the principal author were unsuccessful.

173. Wightman, *supra* note 170, at 12 tbl.1 (reporting means and standard deviations for fall 1991 first-year students from seventy-seven selective law schools).

174. Powers, *supra* note 167, at 740 fig.3, 743 fig.4 (indicating in graphic form that among a sample of twenty-nine law schools with large African American populations, in the vast majority of schools there was considerably greater Black-White overlap on the criterion variable of FYA than there was for admission predictors).

175. Schrader & Pitcher, *supra* note 167, at 488 tbl.9 (converting FYAs and LSATs to standard fifty point scale with a standard deviation of ten). For each of the five law schools in this study the Black-White gap in FYAs was about half of a standard deviation smaller than the LSAT score gaps.

less data is available, the studies by Wightman, Wightman and Muller, and Powers all report a larger Latino-White gap on the LSAT than on FYAs.[176]

Thus, contrary to Subotnik's claim that criticism of the LSAT based on minority performance is "groundless,"[177] the LSAT can be considered racially biased in the fundamental and intuitive sense that the LSAT differentiates Black and Latino law students from White law students to a greater extent than FYAs (or UGPAs). I adopted an unorthodox matching approach to test bias because the predictive validity approach misleadingly lends credence to regressive claims about the fairness of the LSAT.

The studies reviewed above suggest that the LSAT "discriminates" against students of color by artificially magnifying racial/ethnic differences compared to what occurs in college or law school. Given the consistency with which these results occur for such a large number and variety of schools and across time (cohorts from late 1960s to the early 1990s) the greater discriminatory impact of the LSAT is unlikely to result from self-selection patterns within those law schools.[178] It is also unlikely that this general pattern is an artifact of grade inflation or unreliability in first-year law school grades.[179] As with the LSAT, the racial ethnic gap on LSAT/UGPA index scores is larger than the gaps in FYAs (and UGPAs).[180]

---

176. Wightman & Muller, *supra* note 167, at 10 tbl.6b & tbl.6c; Powers, *supra* note 167, at 740 fig.3, 742, 743 fig.4; Wightman, *supra* note 170, at 12 tbl.1.

177. Subotnik, *supra* note 164, at 153.

178. The fact that there are racial differences in admission standards at some law schools (that is, affirmative action) does not sufficiently explain why there is consistently greater overlap between Whites and students of color based on law school grades than LSAT scores (or LSAT/UGPA index scores). In other words, someone contending that selection artifacts produce the smaller gap in FYAs would have to argue that practicing affirmative action would increase the gap in LSAT scores much more than it would correspondingly increase existing gaps in FYAs compared to "race-blind" admissions driven by LSATs and UGPAs. Such a peculiar kind of artifact is dubious. However, "race-blind" admissions might decrease the performance gap in LSATs disproportionately more than the gap in FYAs for non-statistical reasons. Namely, that affirmative action can create a critical mass of ethnic diversity sufficient to prevent alienation and isolation from undermining the law school performance of students of color. *See* Richard O. Lempert et al., *Michigan's Minority Graduates in Practice: Answers to Methodological Queries*, 25 Law & Soc. Inquiry 585, 593 (2000).

> The success of minorities who would have been admitted to Michigan without affirmative action may be due in considerable measure to the existence of the program. Without an affirmative action program, minority students might be present at Michigan in numbers as small as in the post-*Hopwood* Texas Law School that [Professor Thomas] Russell describes. Standing out, as they would if they were present in such small numbers, and without enough fellow ethnics to constitute a community or support group, they might have felt stresses manifested in poorer law school performance.

Id. However, such an environmental explanation properly cannot be considered a statistical artifact of selection procedures. For further discussion of this artifact argument, see David M. White, *Culturally Biased Testing and Predictive Invalidity: Putting Them On the Record*, 14 Harv. C.R.-C.L. L. Rev. 89, 116 n.106 (1979) (criticizing the LSAT for creating greater disparities than occur in law school grades and refuting the notion that factors like self-selection, admission variance, or recruiting efforts could produce this result).

179. The most comprehensive and recent survey of grading practices found that eighty-four percent of the 116 ABA-accredited law schools responding to the survey reported having grade normalization policies. *See* Robert C. Downs & Nancy Levit, *If It Can't Be Lake Woebegone ... A*

2. *Conflating Fairness in Prediction with Fairness in Selection*

A second and related failure of the conventional wisdom is the disjuncture between fairness in regression analysis and fairness in admission policy. This disconnect, along with its adverse consequences, can be demonstrated in a practical context using national law school applicant data. In the 1998-99 admissions cycle to 180 ABA-accredited law schools, there were 47,787 White applicants and 8375 Black applicants, or a ratio of 5.7 White applicants for every Black applicant.[181] Because law schools admit a higher percentage of White applicants overall, the White-Black ratio among admitted applicants is 9.4,[182] which serves as a baseline for how law schools actually admit students in the current environment. If one were developing an admissions policy for a highly competitive law school based upon predictive validity evidence, one could start with UGPA. Indeed, if

---

Nationwide Survey of Law School Grading and Grade Normalization Practices, 65 UMKC L. REV. 819, 836 (1997). Usually, this means that each first-year course must be graded on the same curve as all of the other graded first-year classes. While different schools employ various grading scales, Downs and Levit found that at the forty-eight schools using a 4.0 or 4.33 scale, the average FYA was 2.82. *Id.* at 841.

One of the larger sources of potential unreliability in higher education grades, the noncomparability of grades obtained in different courses, is not so much a problem in law school, at least in the first year. Because law students in a first-year class take essentially all the same courses (with the exception of possibly one Spring elective), efforts to improve the predictive value of law school grades by making the criterion variable of FYAs more reliable have typically failed. *See* John W. Young, *A Comparison of Two Adjustment Methods for Improving the Prediction of Law School Grades*, 55 EDUC. & PSYCHOL. MEASUREMENT 558 (1995) (using item response theory and linear statistical techniques and reporting that they fail to boost the correlation between FYAs and LSAT/UGPA by increasing the reliability of the criterion variable through accounting for differences in courses).

180.   Among students selected and enrolled in law schools there tends not to be much of a positive correlation between LSAT scores and UGPAs (as opposed to a broad group of applicants) when those were the selection criteria. At highly selective schools, the LSAT-UGPA correlation can be negative. *See* Linn & Hastings, *supra* note 81, at 252. They wrote:

Selection on a composite of the LSAT and UGPA tends to produce a lower correlation between the two predictors.... Indeed, following selection, the correlation is often negative. The average correlation for the accepted samples in this study was -.03 and ranged from -.37 to +.34. When both predictors are used in selection, the correlation between the two predictors will decrease from a substantial positive value to a sizeable negative one as the degree of selectivity increases.

Id.; Stephen P. Klein & Roger Bolus, *The Size and Source of Differences in Bar Exam Passing Rates among Racial and Ethnic Groups*, BAR EXAMINER, Nov. 1997, at 8, 15 n.1.

Among the graduates of a law school (as distinct from among those who apply or attend), there is often a very low or even negative correlation between undergraduate grades (UGPAs) and LSAT scores. This stems in part from a school allowing a high LSAT score to offset a low UGPA (and vice versa) in the admissions process.

Id. If LSAT scores and UGPAs have a zero or negative correlation among selected law students, and if students of color perform less well on average than Whites on both predictors, it follows that adding UGPA to the LSAT will not decrease the racial/ethnic gaps compared to the LSAT alone. 1 mention this armchair analysis because in the aforementioned LSAC studies of differential validity and differential prediction, minority-White differences in LSAT/UGPA index scores were not reported.

181.   DECISION PROFILES, *supra* note 125.

182.   *Id.*

our highly selective law school admitted all the national applicants with 3.5+ UGPAs, it would admit 18.2 Whites for every Black applicant (a forty-eight percent drop in Black admits compared to current admission results).[183]

However, in addition to having such a modest correlation to FYAs, UGPA overpredicts African American FYAs by nearly one standard deviation.[184] From a regression perspective, one would want an admissions policy that results in less overprediction (that is, is more predictively "fair" to underrepresented minorities) and greater overall predictive validity. The LSAT satisfies these two requirements. The LSAT correlates about .40 with FYAs nationally[185] and it only overpredicts African Americans' performance by about one-half of a standard deviation.[186] Thus, if our "elite" law school wanted to be more "fair" and admitted all national applicants with LSAT scores of 160 (eighty-third percentile) or higher for example, it would admit 40.1 White applicants for every Black applicant (a seventy-seven percent drop in Black admits compared to current admission results).[187]

In practice, law schools admit applicants largely (though not entirely) on the basis of a combination of LSAT scores and UGPAs,[188] rather than using LSATs or UGPAs separately.[189] Such a practice is certainly in keeping with the conventional wisdom about predictive validity and

---

183. *See* DECISION PROFILES, *supra* note 125. College grades have a quite modest correlation with FYAs of .26 nationally so using them to select law students would select a class that performs better in the first year of law school than a randomly selected class. *See* ANTHONY ET AL., *supra* note 155, at 6.

184. *See* WIGHTMAN & MULLER, *supra* note 167, at 24 tbl.14a (reporting overprediction of 0.8 standard deviations for African Americans); Powers, *supra* note 167, at 746 tbl.7 (reporting overprediction of 1.0 standard deviations for African Americans). It appears that the magnitude of overprediction is going down modestly over time, which is consistent either with the fact that students of color are gradually closing the gap in LSAT/UGPA index scores, or with the fact that since there is a higher percentage of minority law students than in previous deeades, students of color are having a greater influence on the overall regression line.

185. ANTHONY ET AL, *supra* note 155, at 6.

186. *See* WIGHTMAN & MULLER, *supra* note 167, at 24 tbl.14a (reporting overprediction of 0.4 standard deviations for African Americans); Powers, *supra* note 167, at 746 tbl.7 (reporting overprediction of 0.52 standard deviations for African Americans).

187. *See* DECISION PROFILES, *supra* note 125.

188. Indeed, Wightman reports that for the 1990-91 national applicant pool, there was a .78 correlation (sixty-one percent of total variance explained) between LSAT/UGPA combined and actual admission decisions to ABA-accredited law schools for White applicants and a .71 correlation overall. *See* Linda F. Wightman, *An Examination of Sex Differences in LSAT Scores from the Perspective of Social Consequences*, 11 APPLIED MEASUREMENT IN EDUC. 255, 272 (1998); Wightman, *Threat to Diversity*, *supra* note 95, at 11.

189. *See* Kaye, *supra* note 54, at 452 n.89.

> It is important to observe that the entire enterprise of comparing the LSAT's correlation with FYA to that of undergraduate grades may be misdirected. For operational purposes, the more relevant question is whether using the LSAT *in combination with* grades to predict performance is superior to using one of these factors alone. The answer to this question is clearly affirmative.

Id.

differential validity results. Combining LSAT scores and UGPA boosts the correlation with FYAs to about .49 nationally[190] and reduces the overprediction of African Americans to about one-quarter of a standard deviation.[191] Thus, LSAC researcher Donald Powers concludes with respect to the LSAT and UGPA: "[T]he consistent decrease in over- and underprediction resulting from using predictors in combination seems to be a strong reason for using them in this way, according to the regression model definition of fairness."[192] Likewise, LSAC president Philip Shelton rebuts the notion that abandoning the LSAT will increase the number of minorities admitted, arguing that LSAT scores and UGPAs work best when combined as predictors.[193] Referring to her own 1997 LSAC-sponsored national study, Wightman recently concluded that admissions based on a combination of LSAT scores and UGPAs did not substantially disadvantage students of color compared to a UGPA-anchored admissions model:

> Finally, data suggest that test scores are not so much the barriers to admission that many believe them to be. Analysis of law school data investigated the decision outcomes of a 'numbers-only' admission process. The data showed that, regardless of whether the process was modeled by UGPA and LSAT combined or by UGPA only, the consequence would have been a substantial reduction in the overall number of minority applicants who were offered admission to ABA-approved law schools.[194]

Is Wightman's claim, that "numbers-only" admissions based on LSAT/UGPA do not appreciably harm minorities' opportunities compared to "numbers-only" UGPA admissions, supported by the current national applicant data, particularly with respect to highly competitive law schools? Returning to our hypothetical of the highly selective law school, if all applicants with a 3.25+ UGPA and a 155+ LSAT score were accepted, almost the same number of Whites would be accepted as if a 3.5+ UGPA

---

190.   *See* ANTHONY ET AL., *supra* note 155, at 6; Wightman, *Threat to Diversity, supra* note 95, at 32.

191.   *See* Powers, *supra* note 167, at 746 tbl.7 (reporting overprediction of 0.24 standard deviations for African Americans); WIGHTMAN & MULLER, *supra* note 167, at 24 tbl.14a (reporting overprediction of 0.18 standard deviations for African Americans).

192.   Powers, *supra* note 167, at 744; *see also* Wightman, *supra* note 170, at 13 n.29.

> The purpose of each of these studies was to evaluate evidence of the adverse impact of using a common regression system to evaluate the admission applications of black and Hispanic students. Because the data showed that the systems over-predicted subsequent law school performance, each of the studies concluded that using the regression weights did not harm those applicants.

Id.

193.   *See* Shelton, *supra* note 110, at 8.

194.   Linda F. Wightman, *Standardized Testing and Equal Access: A Tutorial, in* COMPELLING INTEREST: EXAMINING THE EVIDENCE ON RACIAL DYNAMICS IN HIGHER EDUCATION ch. 4, 23 (Mitchell   Chang   et   al.,   eds.,   1999)   (online   draft   version),   *available   at* http://www.stanford.edu/~hakuta/RaceInHigherEducation.html (visited Mar. 14, 2001).

alone were required.[195] While combining the LSAT and UGPA might favor an individual Black (or White) applicant by allowing a lower score on one predictor to be compensated by a higher score on the other predictor, the overall result is even worse news for Black applicants. As Figure 1 indicates, requiring both a 3.25+ UGPA and a 155+ LSAT score would admit 54.1 White applicants for every Black applicant (an eighty-seven percent drop in Black admits as compared to current admission results)![196] This data is consistent with the criticism that the current law school meritocracy amounts to a substantial White privilege admissions program.[197]

---

195.   *See* DECISION PROFILES, *supra* note 125. The exact numbers would be 13,456 for 3.5+ UGPAs alone, and 13,248 for 155+ LSATs and 3.25+ UGPAs. Because this data is represented in bands of five points on the LSAT or a quarter-point on UGPA, it is not possible to completely equalize the number of White students before making comparisons (the number of Whites with 160+ LSAT scores was 11,562). Thus, the results of the UGPA, LSAT, and LSAT/UGPA models were expressed as ratios of Whites to Blacks. I realize that my hypothetical is somewhat simplified because a 165/2.75 applicant might have a similar index score as a 155/3.25 applicant, but this fact does not alter the logical structure of the problem I am identifying.

196.   This figure actually may underestimate the hypothetical disparate impact of a combined LSAT/UGPA requirement in the sense that 1 have not weighted LSAT more heavily than UGPA, as LSAC generally recommends based on regression analysis results. Since Blacks fair far worse on the LSAT than on UGPA, placing greater weight on LSAT scores in an LSAT/UGPA index will result in even further disadvantage for African Americans.

197.   *See* Luke Charles Harris & Uma Narayan, *Affirmative Action and the Myth of Preferential Treatment: A Transformative Critique of the Terms of the Affirmative Action Debate*, 11 HARV. BLACKLETTER L.J. 1, 24-25 (1994).

> In a setting where such a wide array of social structures and institutional practices operate to impede equal opportunity for working-class people, women, and minorities, affirmative action policies are not a matter of awarding preferences to members of these groups. Rather, they represent attempts to create greater equality of opportunity than is likely to obtain in the absence of such policies by counteracting some of the effects of the systemic 'preferences' that favor the better-off, white people, and men.

Id.; Daria Roithmayr, *Deconstructing the Distinction between Bias and Merit*, 85 CALIF. L. REV. 1449, 1454-55 (1997).

> First, merit standards necessarily defer to and depend on the very ideas that define social bias and distinguish it from merit. . . . These preferences are necessarily subjective and race-conscious; they are developed in a historically contingent social context and are authored by members of groups who have enough social power—which historically has been based in part on their race and ethnicity—to define what counts as social value.

Id.; Sturm & Guinier, *supra* note 78, at 982. Sturm and Guinier write:

> [C]onventional selection methods advantage candidates from higher socioeconomic backgrounds and disproportionately screen out women and people of color, as well as those in lower-income brackets. When combined with other unstructured screening practices, such as personal connections and alumni preferences, standardized testing creates an arbitrary barrier for many otherwise-qualified candidates.
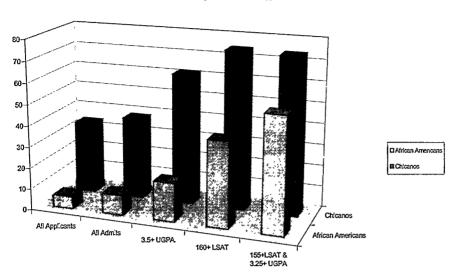
Id.

Figure 1:
Number of White Admits Per Minority Admit, 1998-99 Applicants to ABA Law Schools

While LSAT scores and UGPAs do not entirely determine the admissions process, the above examples illustrate how the pursuit of increasing *fairness* according to the regression model results in greater and greater selection *unfairness* to African Americans as a group. Similar, if less pronounced, results hold for Chicano applicants. In 1998-99 there were 34.0 White applicants for every Chicano applicant, and 38.4 Whites for every Chicano admit.[198] If only 3.5+ UGPA applicants were selected, 61.7 Whites would be admitted for every Chicano admitted (a thirty-eight percent drop in Chicano admits as compared to current admission results).[199] If a 155+ LSAT score and a 3.25+ UGPA were both required for admission, 74.0 Whites would be admitted for every Chicano (a forty-eight percent drop in Chicano admits as compared to current admission results).[200]

The dilemma of "selection bias" highlighted in the aforementioned examples severely burdens students of color in the post-affirmative action landscape. At the highly selective University of Texas Law School, the first three post-*Hopwood* classes (the 2000-02 graduating classes) included only nineteen African Americans out of 1387 students (1.4% of the student body).[201] This is a smaller percentage than in the fall of 1950, when Heman

---

198.   *See* DECISION PROFILES, *supra* note 125.
199.   *See id.*
200.   *See* DECISION PROFILES, *supra* note 125.
201.   Russell, *supra* note 7, at 507.

Sweatt and five other trailblazing African Americans were first permitted to enroll at UT after the Supreme Court declared unconstitutional the law school's mandated policy of de jure segregation.[202] From the perspective of a Black or Latino applying to UT, UCLA, or Boalt Hall, LSAC's claim that the LSAT is "fair" (under the regression model) has a hollow ring. Because there is a larger racial gap on the predictor (LSAT) than on the criterion (FYA), either the LSAT or LSAT testing milieu must be culturally biased,[203] or it must be true that other unmeasured skills and traits that go into being a law student (such as commitment, industry, imagination, motivation, curiosity, and leadership), are significantly less "discriminatory" against underrepresented minorities, since a perfect predictor would perfectly estimate the magnitude of minority-White achievement differences.[204] Psychometricians, beginning with Robert Thorndike, have long recognized the predicament of selection bias.[205] For example, Nancy Cole, current president of ETS, once observed:

> In the regression model . . . the concern is solely with the importance of the criterion at the expense of ideas of fairness to the applicant. However, the potentially successful applicant's concern is often primarily that there be a fair chance of selection regardless of the group membership rather than a guarantee of success. In cases in which the regression line for the group is such that, even though many of that group could succeed, they will have less chance of selection than members of other groups, no selecting institution's concern with selection of applicants with highest predicted criterion scores will receive much sympathy from those applicants. And when poor prediction in one group is the only cause of chances of selection being lowered, the applicant will rightly blame the institution for its failure to find a good predictor—a situation for which the applicant should not be penalized.[206]

Yet, heavy reliance on standardized tests does, in fact, penalize underrepresented minority applicants. Harvard professor Christopher Jencks puts it

---

202.   *Id.* Since the 1950 class was smaller, those six African American students were 2.1% of the entering class. More information about Sweatt v. Painter, 339 U.S. 629 (1950), is available at a Web site set up by Professor Russell. *See* http://www.law.du.edu/russell/lh/sweatt (last visited Mar. 14, 2001).

203.   More specifically, it would be more biased in relative terms if there were also bias disfavoring students of color in undergraduate education and legal education. See *infra* Part IV.C.3 for a discussion of evidence that law school grades are also biased.

204.   *See* Christopher Jencks, *Racial Bias in Testing, in* THE BLACK-WHITE TEST SCORE GAP, *supra* note 20, at 55, 80 ("The conflict between what I will call test-based and performance-based selection systems would vanish if employers could measure all the traits that influence job performance, and could therefore predict exactly how well each applicant would do.").

205.   *See, e.g.,* Robert L. Thorudike, *Concepts of Culture-Fairness,* 8 J. EDUC. MEASUREMENT 63 (1971); Richard B. Darlington, *Another Look at "Culture Fairness",* 8 J. EDUC. MEASUREMENT 71 (1971); Nancy S. Cole, *Bias in Selection,* 10 J. EDUC. MEASUREMENT 237 (1973).

206.   Cole, *supra* note 196, at 253.

more bluntly: "In effect, blacks and Hispanics have to pay for the fact that social science is better at measuring the skills they lack than the skills they have."[207]

Any admissions model, no less one based on linear regression, is unavoidably tied to particular values about what is most important in selecting students.[208] In this light, applying the regression model to law school admissions expresses values that place the prerogatives of White privilege above meaningful measures of fairness for students of color.

3.   The Criterion of First-Year Grades: Too Narrow and Possibly Contaminated by Institutional Racism

Another failure of the conventional predictive validity approach is that choosing FYAs as the criterion to validate the LSAT disadvantages students of color vis-à-vis Whites compared to longer-range criteria, such as second-or third-year grades. For example, in a forthcoming study, Professors Linda Krieger and Marjorie Shultz separately analyzed the first-, second-, and third-year law school performance of students entering Boalt Hall between 1992 and 1996.[209] The authors found that for African Americans, the correlation between LSAT/UGPA index scores and FYA dropped from .50 to .26 to .11 during the three years of law school.[210] In other words, for Black Boalt students, LSAT/UGPA index scores accounted for one quarter of the variance in first-year grades, but a statistically insignificant one percent of variance in third-year grades.[211]

Other research, including some conducted by LSAC, confirms a similar pattern. Powers, in a study of twenty-three ABA-accredited law schools, found that African American and Chicano students made slightly greater increases in second-and third-year law school grades compared to

---

207.   Jencks, *supra* note 204, at 58.

208.   *See* Ronald L. Flaugher, *The Many Definitions of Test Bias*, 33 AM. PSYCHOL. 671, 676 (1978).

> One major advance in our thinking that can be attributed to the development of these various models is that they are now seen to be mathematical expressions of particular value systems, value systems which themselves, potentially at least, can be arrived at by any means whatsoever through the application of whatever criteria are felt to be appropriate whether they be mathematical, ethical, or social.

Id.; *see also* Walter Haney, *Testing and Minorities*, *in* BEYOND SILENCED VOICES 45, 70 (Lois Weis & Michelle Fine eds., 1993); Michael A. Olivas, *Constitutional Criteria: The Social Science and Common Law of Admissions Decisions in Higher Education*, 68 U. COLO. L. REV. 1065, 1081 (1997); Van Zandt, *supra* note 169, at 1495.

209.   Linda Hamilton Krieger & Marjorie Shultz, Validity of LSAT Score & Undergraduate Grade Point Average in Predicting Law Schol Performance Beyond the First Year (unpublished manuscript, on file with author). This study includes an analysis by LSAC done at the request of Boalt Hall, as well as an in-house reanalysis. *Id.*

210.   *Id.* at 17-18 tbl.1. There was also a year-to-year validity drop for Boalt students overall. *Id.* The study included 955 students total, and eighty-nine African Americans. *Id.*

211.   *See id.*

their White classmates,[212] and that for all students the predictive validity of LSAT and LSAT/UGPA index scores declines from the first to the second to the third year of law school.[213] Professor James Hathaway, in an earlier study of Columbia University Law School, concluded:

> The performance of minority students viewed as a group tends to improve significantly vis-à-vis that of their Caucasian counterparts over the course of the J.D. program. . . . The LSAT, which forecasts overall poorer performance by minority students, becomes less and less valid for this group. The net result is that earlier research based solely on first-year statistics may seriously underplay the weakness of the LSAT as a predictor of minority student performance.[214]

In a recent study, Linda Wightman reports that the LSAT is equally effective in predicting law school grades beyond the first year and that there is not a greater drop in prediction among students of color.[215] However, this study relies on cumulative data across three years of law school, and thus does not directly address the issues of differential validity and differential prediction for students of color in the second and third years of law school specifically. Wightman acknowledges the limitation of not having separate second-and third-year data,[216] yet she may not fully

---

212. Donald E. Powers, *Predicting Law School Grades for Minority and Nonminority Students: Beyond the First-Year Average, in* 4 REPORTS OF LSAC SPONSORED RESEARCH: 1978-1983, *supra* note 161, at 261, 287. Powers's sample included 1,200 African Americans and 300 Chicanos. *Id.*

213. Donald E. Powers, *Long-Term Predictive and Construct Validity of Two Traditional Predictors of Law School Performance*, 74 J. EDUC. PSYCHOL. 568, 572 tbl.1 (1982) (reporting on the same twenty-three ABA-accredited schools).

214. Hathaway, *supra* note 172, at 93.

215. LINDA F. WIGHTMAN, LAW SCH. ADMISSION COUNCIL, BEYOND FYA: ANALYSIS OF THE UTILITY OF LSAT SCORES AND UGPA FOR PREDICTING ACADEMIC SUCCESS IN LAW SCHOOL (2000). In this study Wightman concludes:

> The major finding of the study is that LSAT score and UGPA, in combination, were related to cumulative LGPA at approximately the same level as they were related to first-year LGPA. . . . A second important finding is that the patterns of predictive validity for different ethnic and sex groups do not seem to change regardless of whether the criterion is first-year LGPA or cumulative LGPA. . . . The data from the study demonstrate the utility of LSAT scores and UGPAs in the law school admission process beyond the prediction of first-year academic performance in law school, laying to rest a common criticism of their use.

Id. at 2; *see also* Richard O. Lempert et al., *Michigan's Minority Graduates in Practice: The River Runs Through Law School*, 25 LAW & SOC. INQUIRY 395, 459-60 (2000) (citing the aforementioned Wightman study and earlier LSAC research for the proposition that traditional predictors like the LSAT work just as well for cumulative grades). For whatever reason, Lempert's literature review of prediction beyond FYAs made no mention of the Powers or Hathaway studies discussed earlier. *Id.*

216. WIGHTMAN, *supra* note 215, at 37-38. Wightman reasons:

> A third weakness is that the criterion variable cumulative LGPA cannot be disaggregated into first, second, and third year LGPA. This is because the majority of law schools that participated in the *LSAC Bar Passage Study* reported that they do not maintain LGPA separately by year and could not provide the data in that way. . . . Even so, the multiple correlations between the predictors and each of the criteria is so similar that there is nothing in the data to suggest that the results would have been different for this cohort of law school students if the first year LGPA data could have been separated from the grades earned in subsequent years.

appreciate the significance of this limitation. Identical cumulative LGPA validity coefficients can mask racial differences in predictive validity in the second and third years of law school specifically. At Columbia, for example, Hathaway found that even though the correlation coefficients for cumulative LGPA were identical for Whites (.37) and minorities (.38), this masked the fact that coefficients for minorities dropped dramatically from the first (.51) to the second (.27) and third (.17) years of law school, while they were much more consistent for Whites (.35, .33, and .28 respectively).[217]

There is a second reason that Wightman's findings should not, as she suggests, lay to rest a common criticism of reliance on LSAT/UGPA index scores[218] and, by implication, reliance on the predictive validity approach to test bias. This involves the manner in which the LSAT and UGPA are weighted to create index scores that become a principal means of making selection decisions. Usually, law schools weigh the LSAT about 60/40 with grades, a practice that is rooted in the results of LSAC's first-year predictive validity studies.[219] Even if the LSAT and UGPA in combination have the same predictive validity in the second and third years of law school, this does not necessarily mean that the way the two predictors are weighed is optimal in terms of selecting students who will perform best in the latter years of law school. In fact, as Powers's studies suggest, the validity of the LSAT drops while UGPA remains constant.[220] Thus, selection criteria that more heavily weigh LSAT scores will have a disproportionately adverse impact on students of color, and African Americans in particular, because there are greater ethnic differences on the LSAT than in college grades. Similarly, using first-year predictive validity results to justify placing greater weight on the LSAT (versus UGPA) can create a gender bias favoring male candidates, since women have slightly lower average scores on the LSAT and somewhat higher UGPAs.[221]

Id. at 37-38.
   217.   Hathaway, *supra* note 172, at 91 tbl.3.
   218.   WIGHTMAN, *supra* note 215, at 2.
   219.   WIGHTMAN, *supra* note 188, at 270.
   220.   Powers, *supra* note 213, at 572 tbl.1.
   221.   Thus, at ABA-accredited law schools between 1994 and 1998, men had a sixty-nine percent chance of getting admitted to at least one law school, compared to sixty-six percent for women, despite the fact that women have a one-quarter standard deviation advantage in UGPA compared to a one-to two-tenths of a standard deviation disadvantage on the LSAT. Kidder, *Portia Denied, supra* note 47, at 8 n.36, 12 tbl.3. This concern was not raised in Wightman's study, though I believe it has important policy implications. The existence of index-weighting gender bias can also be inferred by studying individual law schools. For example, Boalt Hall is atypical in that it weighs LSATs and UGPAs equally in calculating index scores. I believe that it is no coincidence that the 2000, 1999 and 1998 entering classes at Boalt were 64%, 59% and 55% women. *See* UNIV. OF CAL. AT BERKELEY SCH. OF LAW (BOALT HALL), 2000 ANNUAL ADMISSIONS REPORT tbl.III ("Women Entering Boalt Hall, 1990-2000") (2000). These figures represent a higher proportion of women than at any other "top ten" law school. This information can be obtained by comparing the U.S. News Rankings with the Law School

Yet the problem of using law school grades to answer questions about cultural fairness on the LSAT is much deeper than simply choosing first-year grades instead of second-and third-year grades. Professor Lani Guinier argues that "affirmative action critics' much-touted reliance on *objective measures of merit* have little to recommend them over the life span of a lawyer."[222] When success in the practice of law becomes the benchmark, rather than law school grades, students of color at highly selective institutions, many of whom were recipients of affirmative action, appear to do as well as, and in some cases better than, Whites. For example, Richard Lempert, David Chambers, and Terry Adams conducted an ambitious study of White, Black, Latino and Native American University of Michigan Law School graduates from 1970-96.[223] The authors found that LSAT/UGPA index scores had no correlation with measures of success in the legal profession, such as income or career satisfaction.[224] Not only did minority graduates at Michigan equal their White classmates on these measures,[225] minority graduates made greater contributions to civic service, such as accumulating pro bono hours, serving on the boards of non-profits, and mentoring young attorneys.[226] Bowen and Bok also found that African American graduates of elite colleges who went on to obtain advanced degrees (J.D., M.D., Ph.D., and M.B.A. degrees) were all more likely to lead community service organizations than their White classmates who obtained the same advanced degrees.[227]

Moreover, when psychometricians employ the regression model of test fairness, they are unavoidably adopting a static "black box" approach to validity.[228] Within this framework there is no way to make allowances for the fact that the criterion variable of FYAs can be contaminated by institutional racism. There is a substantial body of literature documenting that various forms of subordination in legal education disproportionately work to the detriment of "outsiders" in law school.[229]

Admission   Council's   link   page   to   ABA   law   schools. http://www.usnews.com/usnews/edu/beyond/gradrank/law/lawtables/gdlaws1.html   and http://www.lsac.org/LSAC.asp?url=lsac/law-school-links.asp (last visited Mar. 16, 2001). *See also* Kidder, *Rise of the Testocracy, supra* note 47, at 179.

222.   Lani Guinier, *Confirmative Action,* 25 LAW & SOC. INQUIRY 565, 582 (2000) (italics added).

223.   Lempert et al., *supra* note 215.

224.   *Id.* at 468 (relying on comprehensive survey data, including a log of income and questionnaires about career satisfaction).

225.   *Id.* at 496-99.

226.   *Id.* at 401.

227.   BOWEN & BOK, *supra* note 63, at 168 fig.6.6.

228.   *See* Mark Kelman, *Concepts of Discrimination in "General Ability" Job Testing,* 104 HARV. L. REV. 1157, 1219 n.175 (1991) (distinguishing between static and dynamic views of validity).

229.   For thoughtful (and non-empirical) discussions of bias against African Americans or Latinos in the structure and environment of legal education, see, for example, Francis Lee Ansley, *Race and the Core Curriculum in Legal Education,* 79 CALIF. L. REV. 1511 (1991); Judith G. Greenberg, *Erasing Race from Legal Education,* 28 MICH. J.L. REFORM 51 (1994); Margaret E. Montoya, *Mascaras,*

For example, Professor Rachel Moran recently completed an intensive interview-based study of Boalt law students in order to assess how Proposition 209 affected Boalt Hall's educational and social climate.[230] In summarizing students' reactions to the first year of law school, Moran concluded:

> In general, then, students perceive the first-year classroom as a hierarchical environment in which their contributions must fit within the professor's agenda. Although students may arrive with diverse experiences and viewpoints, they all must learn to "toe the party line," a process that some find "intellectually stunting." Those students of color who take seriously the notion that their perspectives should be included see themselves in an uphill fight to challenge the status quo. Some believe that they must struggle not only against the professor's unquestioned authority, but also against their classmates' discomfort and anxiety in addressing racial issues.[231]

When learning the law is so enmeshed with learning to "toe the party line," justifying the LSAT based on its correlation with FYAs is dangerous business. I have elsewhere reviewed the legal education literature and found that a wide range of outsider law students—those who are women, minorities, from modest socioeconomic backgrounds, friendly, and dedicated to public interest work—all tend to perform less well than predicted under the regression standard, whereas insider and privileged groups are underpredicted.[232] Ignoring the insider hierarchy of legal education, and strictly applying the conventional regression model of fairness, leads to the following *Alice in Wonderland*-like conclusion: the LSAT is biased against men, Whites, the affluent, those with unsympathetic personalities,

---

Trenzas, y Greñas: Un/masking the Self While Un/braiding Latina Stories and Legal Discourse, 17 HARV. WOMEN's L.J. 185, 201-09 (1994).

230.   Rachel F. Moran, *Diversity and its Discontents: The End of Affirmative Action at Boalt Hall*, 88 CALIF. L. REV. 2241 (2000). This study included in-depth interviews with fifty-nine Boalt students. *Id.* at 2273.

231.   *Id.* at 2283.

232.   Kidder, *Rise of the Testocracy*, *supra* note 47, at 171. *See also* Jomills Henry Braddock II & William T. Trent, *Correlates of Academic Performance among Black Graduate and Professional Students*, *in* COLLEGE IN BLACK AND WHITE: AFRICAN AMERICAN STUDENTS IN PREDOMINANTLY WHITE AND IN HISTORICALLY BLACK PUBLIC UNIVERSITIES 161, 173 (Walter R. Allen et al. eds., 1991) ("For Black professional students, grade performance is explained by a more diverse set of factors including social background factors such as sex and age, major-field competitiveness, interaction with white faculty, and the presence and role of Black faculty in the students' programs."); Pamela Edwards, *The Culture of Success: Improving the Academic Success Opportunities for Multicultural Students in Law School*, 31 NEW ENG. L. REV. 739 *passim* (1997) (reviewing various environmental factors, such as discrimination and social isolation, that might explain the finding that women and students of color perform less well than predicted during the first year of law school).

and those on the corporate law track.[233] Consequently, some psychologists recognize that the notion that overprediction proves the absence of cultural bias on the LSAT is hardly persuasive if overprediction occurs because students of color (or others) encounter substantial biases in the educational setting.[234]

### 4.	The Circularity of Differential Item Functioning

Since the results of this study also could indicate bias in the content of the LSAT, it is necessary to explore how psychometricians typically try to eliminate biased questions on the LSAT. The problems inherent in the conventional approach to item bias shed light on the matching method employed in this study. In addition to predictive validity studies, test producers also seek to ensure the absence of bias on tests like the LSAT by conducting differential item functioning (DIF) analysis of individual questions.[235] DIF is a statistical technique for purportedly identifying specific test items that are disproportionately more difficult for members of a group, such as an ethnic minority, among test takers who have equivalent overall scores on the test.[236]

In the recent *Grutter* trial, the case against the University of Michigan Law School, expert witnesses for the intervening defendants vigorously challenged the efficacy of DIF. *Grutter* experts David White of Testing for the Public, psychology Professor Martin Shapiro of Emory University, and Jay Rosner of the Princeton Review Foundation all criticized DIF's ability to promote fairness on the LSAT.[237] These experts highlighted one of the main problems with DIF as a bias-reducing mechanism: the circular

---

233.	Kidder, *Rise of the Testocracy, supra* note 47, *passim* (using satire to dramatize the negative consequences of interpreting law school performance patterns according to the regression model of fairness).

234.	*See* Robert L. Linn & Charles E. Werts, *Considerations for Studies of Test Bias*, 8 J. EDUC. MEASUREMENT 1, 1 (1971) ("If the prediction equations were found to be identical for all subgroups, this would hardly be evidence of no bias if the criterion itself were biased."); Paul R. Sackett & Steffanie L. Wilk, *Within-Group Norming and Other Forms of Score Adjustment in Preemployment Testing*, 49 AM. PSYCHOLOGIST 929, 932 (1994) (commenting on the regression approach to test bias in the employment context: "First, it is dependent on the quality of the measure of the job-relevant criterion of interest: If that measure is systematically biased against the minority group in question, then the method cannot be used to identify the presence or absence of predictive bias").

235.	James B. Erdmann, *Review of the Law School Admission Test, in* THE THIRTEENTH MENTAL MEASUREMENTS YEARBOOK, 592, 596-97 (James C. Impara & Barbara S. Plake eds., 1998) (summarizing the multiple stages of DIF analysis to which LSAC subjects every LSAT question).

236.	For a description of DIF techniques, see, for example, DIFFERENTIAL ITEM FUNCTIONING (Paul W. Holland & Howard Wainer eds., 1993). For a brief, non-technical description of ETS's DIF procedures, see *What's the DIF? Helping to Ensure Test Question Fairness*, The Educational Testing Service Network, *at* http://www.ets.org/research/dif.html (last visited Feb. 14, 2000).

237.	Expert Report of David M. White, *supra* note 39; Expert Report of Jay Rosner, Grutter v. Bollinger, 16 F. Supp. 2d 797 (E.D. Mich. 1998) (No. 97-75928); Expert Report of Martin M. Shapiro, Grutter v. Bollinger, 16 F. Supp. 2d 797 (E.D. Mich. 1998) (No 97-75928). 1 should disclose that 1 provided research support for David White's report, including in the section on DIF. All three experts testified in the *Grutter* trial in early 2001.

reasoning involved in controlling for total score on the test before looking for items that are differentially difficult.[238] As psychometrician Lorrie Shepard and others have noted, controlling for total test score means that there is no external standard for judging what is or is not fair.[239] The absence of an external standard of bias means that DIF techniques are incapable of rooting out systemic bias against a minority group;[240] by controlling for total score, it is a foregone conclusion that questions biased "against" a group must be balanced out by questions "in favor" of that group.[241]

While researchers for the major test producers have acknowledged DIF's tautological limitation,[242] little has been done to remedy the situation. Test developers have almost universally opposed one suggested alternative: the "Golden Rule" technique. With this technique, questions are

---

238.   Lorrie Shepard et al., *Comparison of Procedures for Detecting Test-Item Bias with Both Internal and External Ability Criteria*, 6 J. EDUC. STAT. 317, 321 (1981).

> A major limitation of all of the bias detection approaches employed in the research to date is that they are all based on a criterion internal to the test in question. They cannot escape the circularity inherent in using total score on the test or the average item to identify individuals of equal ability and hence specify the standard of unbiasedness.

Id.

239.   *Id.*

240.   William H. Angoff, *Perspectives on Differential Item Functioning Methodology*, in DIFFERENTIAL ITEM FUNCTIONING, *supra* note 236, at 3, 17 ("For if the criterion is itself biased to some degree, then the application of a DIF analysis will certainly be flawed; further, if bias is pervasive in the criterion, then any attempt to identify bias in its component items will inevitably fail."); Kidder, *Portia Denied*, *supra* note 47, at 35.

> If gender or racially biased questions have spill-over consequences affecting performance on subsequent questions, or if stereotype threat creates a constant bias permeating the entire test (or both), the LSAC's bias detection methods would reveal no bias precisely when bias was most severe and systematic.

Id.

241.   Gregory Camilli, *The Case against Item Bias Detection Techniques Based on Internal Criteria*, in DIFFERENTIAL ITEM FUNCTIONING, *supra* note 236, at 397, 409.

> The circulatory problem is exemplified by the fact that the average DIF index on a test is zero. Internal methods of DIF are ipsative: Holding ability constant, if one group of examinees tends to miss some items unexpectedly, it must unexpectedly answer other items correctly. In other words, items that disfavor the minority group are canceled by items that favor the minority group.

Id. at 409.

242.   Nancy S. Cole, *Judging Test Use for Fairness*, in U.S. COMM'N ON CIVIL RIGHTS, THE VALIDITY OF TESTING IN EDUCATION AND EMPLOYMENT 92, 102 (1993) (Cole, currently the President of ETS, acknowledged that DIF cannot "guarantee that there is no gender bias in the questions."); Howard Wainer, *Precision and Differential Item Functioning on a Testlet-Based Test: The 1991 Law School Admissions Test as an Example*, 8 APPLIED MEASUREMENT IN EDUC. 157, 182 (1995) (conducting an ETS-sponsored study of DIF analysis on the LSAT). Wainer noted:

> The size of these differences in the worst case was small enough to suggest that no serious problems of fairness exist. Moreover, even these small differences were reduced to almost nothing through the balancing across testlets. This balance cannot be swallowed whole. Because performance on the test section itself determined the stratifying variable, the overall balance (zero overall DIF) is almost tautological. That the balancing works as well as it does at all levels of examinee proficiency is not mathematically determined.

Id.

categorized based on the magnitude of subgroup differences and items with the largest gaps are excluded from the final version of the test.[243]

Rutgers psychometrician Gregory Camilli suggests another alternative: conduct DIF analysis of LSAT items using external criteria rather than total score on the test.[244] Similarly, in his report in the *Grutter* case, David M. White concludes that when using DIF techniques "a different comparison would occur if subgroups were compared on the basis of other measures of accomplishment, for example, by their performance in college, as measured by GPA."[245]

As a practical matter, however, LSAC and other test developers have failed to use external criteria like UGPA to validate the fairness of the LSAT. For instance, in 1990, LSAC researchers Linda Wightman and David Muller used traditional DIF techniques on the LSAT, and found (quite unsurprisingly, given the circularity problem discussed earlier) that the LSAT items were not biased against students of color or women.[246] Wightman and Muller recognized the insufficiency of exclusive reliance on internal test criteria, however, and thus recommended, "Further work to analyze minority test performance should match test takers on some criterion, such as undergraduate grade-point average . . . before comparing performance."[247] In the decade since this recommendation, however, no LSAC-sponsored research has keyed item-fairness measurements to the

---

243. The "Golden Rule" method refers to a 1984 settlement between ETS and the Golden Rule Insurance Company over alleged racial bias on the Illinois Insurance Exam produced by ETS. This led to a similar settlement in an Alabama teacher exam, and to proposed legislation in several states. By 1987, ETS President Gregory Anrig recanted the Golden Rule settlement and declared that ETS had made a mistake. For background information and contrasting views on the Golden Rule settlement as well as the legislation it spawned, see Lloyd Bond, *The Golden Rule Settlement: A Minority Perspective*, EDUC. MEASUREMENT ISSUES & PRACTICE, Summer 1987, at 18; Jane Faggen, *Golden Rule Revisited: Introduction*, EDUC. MEASUREMENT ISSUES & PRACTICE, Summer 1987, at 5; Richard M. Jaeger, *NCME Opposition to Proposed Golden Rule Legislation*, EDUC. MEASUREMENT ISSUES & PRACTICE, Summer 1987, at 21; Robert L. Linn & Fritz Drasgow, *Implications of the Golden Rule Settlement for Test Construction*, EDUC. MEASUREMENT ISSUES & PRACTICE, Summer 1987, at 13; J. Patrick Rooney, *Golden Rule on Golden Rule*, EDUC. MEASUREMENT ISSUES & PRACTICE, Summer 1987, at 9; John Weiss, *The Golden Rule Bias Reduction Principle: A Practical Reform*, EDUC. MEASUREMENT ISSUES & PRACTICE, Summer 1987, at 23.

244. Camilli, *supra* note 241, at 409 ("If DIF is linked to test bias, it is not by way of indicating a systematic underestimation of test scores for a particular group. It must be linked by an unbiased external criterion, or by expert judgment.").

245. Expert Report of David M. White, *supra* note 45, at 13.

246. *See* LINDA F. WIGHTMAN & DAVID G. MULLER, COMPARISON OF LSAT PERFORMANCE AMONG SELECTED SUBGROUPS, LAW SCHOOL ADMISSION COUNCIL STATISTICAL REPORT 90-01, 29 (1990).
> More importantly, examination of data for evidence of differential difficulty for some items, a possible indicator of biased or unfair items, does not identify problem items. . . .
>
> . . .
>
> . . . There is no evidence that any one item type particularly disadvantages minority test takers nor are there individual items that exhibit statistical evidence of bias toward any subgroup."

Id.

247. *Id.* at 29.

external criterion of UGPA. Similarly, in a 1995 ETS-sponsored study of DIF on the LSAT, Howard Wainer recommended that future work should use law school grades as a DIF criteria.[248] This, too, has yet to be conducted.

One way to conceptualize this study is as an attempt to realize Wightman and Muller's suggestion of matching LSAT test takers on UGPA. What distinguishes this effort is a much larger unit of analysis (total test score, rather than performance on individual items)[249] and rigorous controls governing the matching of applicants with equivalent UGPAs (such as institution attended, graduation date, and choice of major). The present results and discussion show how the LSAT creates greater levels of racial stratification than either UGPA or FYA or success in legal practice.

D.   Evaluating Likely Alternative Hypotheses

1.   Does Grade Compression Compromise UGPA as a Matching Criterion?

One possible problem in this study design is grade inflation, which could affect the results in two ways. First, within the same college, standards can change over time. This issue was addressed in the initial study design by controlling for graduation date. Second, grade inflation can compress the grading scale and undermine the reliability of UGPA, which is the degree to which a standard is consistently free of measurement error. In other words, if colleges issue mostly "A's" and "B's," there would be a stacking effect at the upper end of the UGPA distribution, making it difficult to differentiate the exceptional student from the very good student.[250] Such bunching would make UGPA a problematic matching criteria, or so the argument goes.[251] Thus, before concluding that the LSAT is racially biased, it is important to assess whether UGPA unreliability due to grade compression is likely to exaggerate LSAT score differences by race/ethnicity among law school applicants that have been matched on UGPA within the same college.

For law school admission purposes, UGPA reliability has not been generally studied as extensively as LSAT reliability. Nonetheless, available

---

248.   Wainer, *supra* note 242, at 183 ("One prospective study would be to see how predictive of success in law school are each of the passages. One can analyze such data in a way faithful to the structure of the test by using law school grades as the stratifying variable in a DIF study.").

249.   Presumably, total test score is a more effective unit of analysis for capturing the effects of consistent or pervasive bias on LSAT questions.

250.   *Cf.* Randal C. Archibold, *Princeton Re-Evaluates Grading Standards*, N.Y TIMES, Feb. 12, 1998, at B12 (reporting that between 1992 and 1997, eighty-three percent of the grades awarded at Princeton were between a B- and an A+).

251.   This criticism was raised in response to Testing for the Public's presentation to the Texas Legislature. Dewar & Kamath, *supra* note 93, at 1 (quoting both David Murray of the Statistical Assessment Service and Harvard Professor Stephan Thernstrom that grade inflation at elite schools makes UGPA an unreliable matching criterion).

estimates from LSAC place the reliability coefficient of cumulative UGPA at .92,[252] which is in the same range as the .90-.95 reliability coefficients reported for the LSAT.[253] The most extensive research on college grades and the reliability of those grades was conducted by Warren Willingham and his colleagues at ETS and the College Board.[254] In summarizing the evidence based on ETS's exhaustive data sources, Willingham concluded:

> These data belie a common misconception: that grade inflation will necessarily make the GPA less predictable. From a technical standpoint, there is actually little reason to expect that moderate overall shifts in the grading level, such as have been observed over the years, would have much effect on validity coefficients.[255]

Willingham's finding requires explanation, particularly since the belief that grade inflation necessarily produces UGPA unreliability dies hard.[256] If grading at a college is "inflated," so that mostly "A's" and "B's" are issued, the rating scale will still have about six points of differentiation when straight letter grades are modified by plus and minus marks.[257] In general, the reliability of a rating scale tends to be relatively constant if the scale includes five to seven points.[258] Thus, it is not surprising that Willingham's finding that grade compression typically fails to decrease the reliability of UGPA is consistent with Bejar and Blew's meta-analysis for the College Board,[259] with Millman's research on Cornell undergraduates,[260] and with Singleton and Smith's study of UC Riverside students.[261]

---

252. Alfred B. Carlson & Charles E. Werts, *Relationships among Law School Predictors, Law School Performance, and Bar Examination Results, in* 3 REPORTS OF LSAC SPONSORED RESEARCH: 1975-1977, *supra* note 155, at 211, 220; *cf.* LAW SCHOOL ADMISSION REFERENCE MANUAL, *supra* note 126, at 3.5 (reporting that "undergraduate records, being the accumulation of four years' grades, tend to have high reliability").

253. LAW SCHOOL ADMISSION REFERENCE MANUAL, *supra* note 126, at 3.5.

254. WARREN W. WILLINGHAM ET AL., PREDICTING COLLEGE GRADES: AN ANALYSIS OF INSTITUTIONAL TRENDS OVER TWO DECADES (1990).

255. *Id.* at 71 (citation omitted). Willingham reports that college grades became inflated at the same time that SAT-UGPA validity coefficients went up, and that during the period when SAT-UGPA validity coefficients came back down, there was no indication of grade inflation. *Id.* Both of these findings confound the thesis that grade inflation necessarily lowers the reliability of UGPA, though neither result necessarily proves that grade inflation cannot lower UGPA reliability.

256. *See* Diane F. Halpern, *Validity, Fairness, and Group Differences: Tough Questions for Selection Testing,* 6 PSYCHOL., PUB. POL'Y, & L. 56, 58 (2000) ("For many purposes, college grades serve as the criterion against which admissions tests are validated, but grades are unreliable and grade inflation is making them increasingly worthless for any purpose.").

257. Under LSAC's grading standardization process, these points would be A+ (4.33), A (4.0), A- (3.67), B+ (3.33), B (3.0), B- (2.67) and C+ (2.33). *See* LAW SCH. ADMISSION COUNCIL, *supra* note 73, at 22.

258. Royce Singleton, Jr. & Eliot R. Smith, *Does Grade Inflation Decrease the Reliability of Grades?,* 15 J. EDUC. MEASUREMENT 37, 37-38 (1978).

259. Isaac I. Bejar & Edwin O. Blew, *Grade Inflation and the Validity of the Scholastic Aptitude Test,* 18 AM. EDUC. RES. J. 143, 155 (1981). Bejar and Blew write:

> If grade inflation is viewed as the addition of a constant to grades it will not affect validity because the addition of such a constant would not affect the correlation of GPA and SAT. Beyond a point, however, the variability of GPA might be restricted causing a reduction in

2001] RACIAL AND ETHNIC DIFFERENCES AND THE LSAT 1111

Analyzing a subset of the Boalt applicant database provides an independent source of support for the conclusion that grade compression-related unreliability is not spuriously producing the racial LSAT gaps among academic peers. The grades issued at UC Berkeley are less compressed than those issued at elite private universities like Stanford and Princeton. In 1996-98, the average UC Berkeley UGPA was 3.10,[262] compared to 3.4 for Stanford and Princeton.[263] In addition, only one quarter of UC Berkeley students obtained UGPAs of 3.50 or higher during this time.[264] Similarly, the average UGPA for all 1996 UC Berkeley applicants to ABA-accredited law schools was 3.21.[265] Since UC Berkeley is the single largest source of information in the database, it is possible to present separate matching results for UC Berkeley alone. If grade compression were a major source of the racial and ethnic LSAT differences among students matched on UGPA within the same college, one would expect that the LSAT performance differences would be considerably smaller at UC Berkeley than at other more grade-inflated universities.

The data does not support this grade inflation hypothesis. For UC Berkeley, considered separately, the LSAT gaps were slightly larger than the overall gaps for African Americans and Latino applicants, and the gap was basically unchanged for Asian Pacific Americans.[266] In summary, this evidence from UC Berkeley, along with the previously discussed evidence about grade compression and reliability from ETS and other sources,

the correlation of GPA and SAT. The data presented do not suggest that point has been reached.

Id.

260.   Jason Millman et al., *Does Grade Inflation Affect the Reliability of Grades?*, 19 RES. IN HIGHER EDUC. 423, 428 (1983) ("When plus and minus grades are possible, even with grade inflation there remains sufficient opportunity to discriminate among students reliably.").

261.   Singleton & Smith, *supra* note 258, at 41 ("However, our data show that grade inflation does not necessarily entail a decrease in grading reliability or its concomitants, a lack of discrimination between students with different levels of true ability and lowered predictive power.").

262.   UNIV. OF CAL., BERKELEY OFFICE OF STUDENT RESEARCH, BERKELEY UNDERGRADUATE FACT SHEET—FALL 1988 (1999) (listing the overall UGPA for UC Berkeley students in 1998 as 3.10); UNIV. OF CAL., BERKELEY OFFICE OF STUDENT RESEARCH, BERKELEY UNDERGRADUATE FACT SHEET—FALL 1996 (1997) (listing the overall UGPA for UC Berkeley students in 1996 as 3.08).

263.   Archibald, *supra* note 70 (reporting that the median UGPA at Princeton in 1997 at 3.4); Gose, *Efforts to Curb Grade Inflation Get an F from Many Crities, supra* note 70, at A41 (reporting that the average UGPA at UC Berkeley in the Fall 1996 was 3.10, compared to 3.3 at Duke and 3.28 at Dartmouth).

264.   UNIV. OF CAL., BERKELEY OFFICE OF STUDENT RESEARCH, BERKELEY UNDERGRADUATE FACT SHEET—FALL 1996 (1997).

265.   LAW SCH. ADMISSION COUNCIL, MASTER LAW SERVICES REPORT FOR WILLIAM C. KIDDER (1997). Master Law Services Reports are sent to each law school applicant who registers with Law Services and include summary information about the applicant's college grades and LSAT scores. Since I applied to law school in the fall of 1997 as a UC Berkeley graduate, my Master Law Services Report included information about the mean UGPA of all UC Berkeley law school applicants from the previous year.

266.   The LSAT gaps for UC Berkeley were as follows: African Americans 10.3 points (n = 82), Latinos 7.4 points (n = 148), and Asian Pacific Americans 2.6 points (n = 384).

should put to rest the notion that the LSAT gaps identified in this study are an artifact of a compressed, unreliable matching criterion.

2. *Are College Grades Biased in Favor of Students of Color?*

Some of the more strident affirmative action critics, such as Michigan Law Professor Terrance Sandalow,[267] Center for Equal Opportunity Litigation Director Roger Clegg,[268] Harvard Professor Stephan Thernstrom,[269] Manhattan Institute scholar Abigail Thernstrom,[270] and others[271] argue that college grades are biased in favor of minority students because liberal guilt dictates grading leniency. Adherents of such a view can be expected to raise objections to the present matching study based on this hypothesis. Are the views of Sandalow and others supported by empirical studies of higher education? Responding to Sandalow's critique of *The Shape of the River*, Bowen and Bok argue that Sandalow lacks data to support his inflammatory view, and findings from the College and Beyond database suggest a contrary position.[272]

In fact, there is more empirical support for the proposition that institutional racism influences performance in college and disproportionately harms students of color.[273] Colette van Laar and her colleagues, for

---

267.   Terrance Sandalow, *Minority Preferences Reconsidered*, 97 MICH. L. REV. 1874, 1903 (1999) (book review). In a review of *The Shape of the River*, Sandalow argues:

> Whatever the reasons, the reluctance of faculties to award grades that would lead to academic dismissals or to voluntary decisions to withdraw because of discouragement has necessarily had an effect on decisions about the quality of work that is to be regarded as minimally acceptable. . . . 'Grade inflation' is not, of course, solely attributable to minority admission policies, but they are surely one of the factors that have contributed to it. The overall effect is a lowering of academic standards . . . from those that might reasonably be expected currently at the nation's premier academic institutions.

Id.

268.   Roger Clegg, *Why I'm Sick of the Praise of Diversity on Campuses*, CHRON. HIGHER EDUC., July 14, 2000, at B8 ("If colleges and universities lower standards to achieve diversity in admissions, diversity advocates will inevitably pressure those institutions to rig the requirements in grading and for graduation.").

269.   Stephan Thernstrom & Abigail Thernstrom, Letters from Readers, 107 COMMENTARY 5, 20 (May, 1999) (noting that "the troubling issue of 'race conscious' grading . . . squares with anecdotal evidence we have seen").

270.   *Id.*

271.   Paul Hollander, Letters from Readers, 107 COMMENTARY 5 (May, 1999). A University of Massachusetts professor responded to a Thernstrom piece by arguing:

> Though hard to prove or quantify, 'race-conscious' grading (that is, going easy on black students) is a fact of life, especially in the humanities, social sciences, and schools of education. It is very hard to know how many more black students would drop out if this were not the case. Such grading is usually motivated by compassion or sympathy or fear of being called a 'racist' if black students were to do poorly in one's course.

Id. at 19.

272.   William G. Bowen & Derek Bok, *Response to Review by Terrance Sandalow*, 97 MICH. L. REV. 1917, 1921-22 (1999) (questioning Sandalow's argument by citing Black-White similarities in percentages majoring in engineering, mathematics, chemistry, and biology); *see also* Gannon, *supra* note 5, at 281 (critiquing the grading leniency argument).

273.   This fact is not inconsistent with the finding that UGPA has high reliability.

example, studied a random sample of about 5700 undergraduates at UT-Austin,[274] one of the feeder schools in the present study. Students who were more racist in their attitudes obtained better college grades in majors that tended to cultivate "hierarchy-enhancing" social values,[275] compared to less racist students in the same majors.[276]

In a national survey, UCLA sociologist Walter Allen compared the performance of African Americans in predominantly White and historically Black public colleges.[277] Allen reported that African American students attending historically Black colleges had better academic performance, social activism, and career aspirations.[278] Allen found that Black students on predominantly White campuses emphasized feelings of alienation and racial discrimination, whereas those on historically Black campuses emphasized having a sense of engagement, connection, and encouragement.[279] Allen concluded, "Finally, little doubt exists over the negative impact of hostile racial and social relationships on Black student achievement. When Black students are made to feel unwelcome, incompetent, ostracized, demeaned, and assaulted, their academic confidence and performance understandably suffer."[280]

These findings about campus climate are not atypical.[281] A study conducted on behalf of student of color organizations intervening in *Grutter*, the University of Michigan Law School affirmative action case, is relevant to this study of elite colleges. With financial support from the Society of American Law Teachers (SALT), the intervenors had Walter Allen and Daniel Solorzano conduct a campus climate study, including surveys and focus groups, of several feeder schools to the University of Michigan Law

---

274.   Colette van Laar et al., *The Three Rs of Academic Achievement: Reading 'Riting, and Racism*, 25 PERSONALITY & SOC. PSYCHOL. BULL. 139 (1999).

275.   "HE belief systems are defined as those that provide moral or intellectual justification for the establishment or maintenance of hierarchical and antiegalitarian relations among social groups. Examples of HE belief systems would be ideologies such as classism, sexism, aggressive nationalism, and racism." *Id.* at 140.

276.   *Id.* at 147. The van Laar study tracks institutional rewards for racism, which is part of the same larger dynamic by which societal racism places obstacles that can adversely affect the performance of students of color. The "hierarchy-enhancing" majors included economics, accounting, finance, general business, marketing, management, and advertising. *Id.* at 145.

277.   Walter R. Allen, *The Color of Success: African-American College Student Outcomes at Predominantly Black Public Colleges and Universities*, 62 HARV. EDUC. REV. 26 (1992).

278.   *Id.* at 35-37.

279.   *Id.* at 39.

280.   *Id.* at 41.

281.   *See, e.g.*, Joe R. Feagin & Melvin P. Sikes, *How Black Students Cope with Racism on White Campuses*, 8 J. BLACKS HIGHER EDUC. 91 *passim* (1995); Sylvia Hurtado et al., *Enhancing Campus Climates for Racial/Ethnic Diversity: Educational Policy and Practice*, 21 REV. HIGHER EDUC. 279 *passim* (1998). For a review of the recent social psychology literature on racism in higher education, see Shana Levin, *Social Psychological Evidence on Race and Racism, in* COMPELLING INTEREST: EXAMINING THE EVIDENCE ON RACIAL DYNAMICS IN HIGHER EDUCATION, *supra* note 194, at 1-17.

School, including Michigan, Michigan State, Harvard, UC Berkeley, and UCLA.[282] Recall that in the Boalt applicant database, UC Berkeley, UCLA, and Harvard are three of the top four feeder schools to Boalt, and Michigan was also in the group of fifteen schools.[283] Overall, Allen and Solorzano found that White privilege and entitlement are overarching features of the undergraduate racial climate, and that the cumulative effect of the educational environment created an uneven playing field for students of color.[284] Thus, it should be remembered that this study of the LSAT provides insight into the scope of bias on the LSAT relative to that on other measures like UGPA and FYA.

3. Why Should LSAT Scores Be Similar in the Absence of Bias?

Another objection raised by LSAC officials in response to a presentation of the preliminary results of this study is that equivalent performance in college should not be expected to correspond to equivalent performance on the LSAT in the first place.[285] Beneath this contention is the notion that the LSAT is carefully designed to measure skills associated with first-year law school performance, whereas undergraduate grades reflect mastery of college coursework, which is subject to numerous random influences that distinguish its content domain from that involved in first-year law school grades. There are, however, three reasons why we should expect UGPA to correlate closely to LSAT scores. First, LSAC's own research indicates that such a correlation should exist. Second, there is no reason to believe that the LSAT tests for specialized legal analysis skills unrelated to undergraduate success. Finally, law schools organize their admissions policy around the assumption that these two measures correlate.

The first problem with the claim that UGPAs should not be expected to correlate with LSAT scores is that it contradicts LSAC's own research. LSAC has found that LSAT scores and UGPAs correlate about .40 with each other in the national applicant pool.[286] While this position also

---

282. Expert Report of Walter Allen & Daniel Solorzano, Grutter v. Bollinger, 16 F. Supp. 2d 797 (E.D. Mich. 1998) (No. 97-7592), *available at* http://www.bamn.com/literature/lit-um-case-expert-rpts.htm. (visited Feb. 18, 2001). This report will be republished in 12 LA RAZA L.J. (forthcoming 2001).

283. *See infra* Part II.A.

284. Expert Report of Allen & Solorzano, *supra* note 282, at 56-59.

285. *See* Dewar & Kamath, *supra* note 93, at 1 (quoting LSAC Associate Corporate Counsel James Vaseleck, who stated, "the faulty assumption in the study's analysis is that students with the same GPA from the same school should have the same test scores"). Vaseleck also contended that the LSAT is a test of skills which may or may not be attained through mastery of college coursework. *Id.*; *see also* Chris Jenkins, *Study Highlights Disparities in LSAT Scores*, DAILY CALIFORNIAN, Oct. 29, 1998, at 1 ("The assumption of the study that LSAT and GPA scores measure the same thing is wrong.") (quoting Vaseleck).

286. Linn & Hastings, *supra* note 88, at 252 (noting that the correlation between UGPAs and LSATs can be expected to be about .40 for an unselected population of law school applicants); Wightman, *Threat to Diversity*, *supra* note 95, at 31 (noting that "the LSAT score and UGPA of law

contradicts the advice LSAC officials give to minority applicants preparing for the LSAT,[287] the more serious objection pertains to the historical evidence about the LSAT itself.

Whether or not the LSAT uniquely captures the skill dimensions required for law school performance can be assessed by studying the relationship between the LSAT and FYA over time. The LSAT has undergone major content and format changes in the last twenty years with negligible impact on the predictive validity of the test. For example, in 1978 the LSAT consisted of the following sections: principles and cases, error recognition, sentence correction, logical reasoning, practical judgment, and quantitative comparison (math).[288] LSAC's national validity study for the next year estimates the correlation coefficient between the LSAT and FYA at .40.[289] The LSAT was changed in 1982,[290] and when changed again in 1989, three scored sections remained: analytical reasoning, reading comprehension, and logical reasoning.[291] For the 1990 entering class, the correlation between the LSAT and FYA was .41.[292] After 1991, a second section of logical reasoning was added and each section was shortened in length.[293] The latest validity study reports that the average correlation

---

school applicants are correlated .38 with one another. Thus, there is some redundancy in these measures."); *see also* Philip Shelton, *Executive Director's Report: Top Ten Misconceptions about the LSAT*, LAW SERVICES REP. 4 (Jan./Feb. 1999) (reporting that the relationship between the LSAT and UGPA is "fairly strong"); WIGHTMAN, *supra* note 155, at 16 (stating that "In a random group of applicants, this [LSAT and UGPA] correlation would be fairly high, indicating that applicants with high LSAT scores also had high UGPAs, while applicants with low LSAT scores also had low UGPAs."). While a .40 correlation may be nothing to write home about, it is sufficiently high that one might expect applicants with the same UGPAs (especially under the controlled conditions in this study) to do approximately the same on the LSAT. After all, the core justification for using the LSAT is its correlation with FYAs, and this correlation is also quite modest. See *infra* Part IV.C. for a discussion of the LSAT as a predictor.

287.   *See* LAW SCH. ADMISSION COUNCIL, THE TEN QUESTIONS MOST FREQUENTLY ASKED BY MINORITY APPLICANTS TO LAW SCHOOL (1997) (stating, in response to the question, "How should I prepare for the LSAT?," "The best preparation for the test is a solid undergraduate education with an emphasis on reading and reasoning skills, linked to familiarity with the essentials of the test."); *see also* Kent Lollis, Assistant to the President, Presentation at the Minority Affairs at LSAC Race Judicata Conference, University of New Mexico School of Law, Panel on the LSAT (Mar. 9, 2000) (videotape on file with author) (stating that minorities' performance on the LSAT could be significantly improved by taking particular courses in college).

288.   LAW SCH. ADMISSION COUNCIL, *supra* note 126, at 3 fig.1 (listing the component sections of the LSAT at various points in time).

289.   Franklin R. Evans, *Recent Trends in Law School Validity Studies*, *in* 4 REPORTS OF LSAC SPONSORED RESEARCH: 1977-1983, *supra* note 161, at 347, 355 tbl.3.

290.   The Data Evaluation and Application section was removed in 1982, possibly because it was easily amenable to coaching. *See* Brain Powell & Lala Carr Steelman, *Equity and the LSAT*, 53 HARV. EDUC. REV. 32, 40-41 (1983) (providing experimental evidence demonstrating that only two hours of coaching on the Data Evaluation and Application section produced overall test score gains equivalent to thirty-three points on the 200-800 scale compared to a control group).

291.   LAW SCH. ADMISSION COUNCIL, *supra* note 126, at 4.

292.   LINDA F. WIGHTMAN, LAW SCH. ADMISSION COUNCIL, PREDICTIVE VALIDITY OF THE LSAT: A NATIONAL SUMMARY OF THE 1990-1992 CORRELATION STUDIES 10 tbl.2 (1993).

293.   LAW SCH. ADMISSION COUNCIL, *supra* note 126, at 4.

between the LSAT and FYA was .40 for the entering classes of 1995 and 1996.[294] As reported earlier, these correlations are of the same magnitude as the correlation between the LSAT and UGPA for the national pool of law school applicants.[295]

Thus, whether the LSAT includes sections like sentence completion, math, reading comprehension, or principles and cases appears to have little if any influence on the magnitude of its correlation with FYAs, which consistently hovers around .40. This level of correlation casts doubt on the notion that the LSAT is somehow uniquely designed to capture the set of skills required for the study of law. Rather, the data suggest that any major norm-referenced "aptitude" test could stand in for the LSAT and produce equivalent results.[296] Indeed, earlier LSAC research found that the Graduate Record Exam (GRE) subject knowledge tests (predominantly taken in business, economics, government, history, and literature) were a better predictor of FYAs than the LSAT,[297] despite the fact that GRE domain-specific knowledge tests were never designed to predict law school performance. Furthermore, the correlations among tests like the SAT, LSAT, and GRE usually account for three times as much variance as the correlations between these tests and what they are intended to predict (college, law school, and graduate school grades, respectively).[298] Collectively,

294.   ANTHONY ET AL., *supra* note 155, at 6 tbl.2.

295.   To be fair, this is comparing a restricted correlation to an unrestricted correlation.

296.   *Cf.* Gary B. Melton, *Review of Law School Admission Test, in* 1 THE NINTH MENTAL MEASUREMENTS YEARBOOK 824, 825 (James V. Mitchell, Jr. ed., 1985). Melton reports that sections of the LSAT with the greatest facial validity, such as Issues and Facts, actually have the lowest predictive validity, and concludes:

> This point raises the question, apparently still unanswered, posed in the two previous [*Mental Measurements Yearbook*] reviews of the LSAT: is a separate law aptitude test really needed? Although 'thinking like a lawyer' may be significantly different from the reasoning skills nurtured in graduate studies, it is not at all self-evident that aptitude for the former cannot be predicted by the Graduate Record Examination as well as, or better than, the LSAT. This hypothesis deserves empirical investigation.

Id.

297.   W.B. Schrader & Barbara Pitcher, *The Advanced Tests of the Graduate Record Examinations as a Predictor of Law School Grades, in* 2 REPORTS OF LSAC SPONSORED RESEARCH: 1970-1974, *supra* note 167, at 47, 48 ("GRE Advanced Test scores tended to show higher correlation with law school grades than did LSAT scores."). To the best of my knowledge, there has not been another LSAC-sponsored study of the predictive validity of other non-LSAT standardized tests.

298.   ALEXANDER ASTIN, WHAT MATTERS IN COLLEGE 213 (1993) (reporting, for a nationally representative sample of 5800 law school candidates, that the correlation between the LSAT and the SAT was .72); ALLAN NAIRN & ASSOCS., THE REIGN OF ETS: THE CORPORATION THAT MAKES UP MINDS 234 (1980) (commenting on the Schrader study, and reporting that ETS and LSAC found that the correlation between the LSAT and SAT was so high that "the feeling in the Council was that undergraduate advisors were not generally equipped to make appropriate use of such information"); William H. Angoff & Eugene G. Johnson, *The Differential Impact of Curriculum on Aptitude Test Scores*, 27 J. EDUC. MEASUREMENT 291, 294 (1990) (reporting for a national sample of about 23,000 that the SAT/GRE verbal sections correlate with each other .86, and the SAT/GRE quantitative section have a .86 correlation); W.B. Schrader, *Score Relations Study, in* 1 REPORTS OF LSAC SPONSORED RESEARCH: 1949-1969, 247, 254 tbl.3 (Law Sch. Admission Council ed., 1976) (providing a table of what ranges of scores on the LSAT students could expect based on their SAT scores); Expert Report of

these findings suggest that nothing even approaches being a "pure" test of law school aptitude, which, in turn, bolsters the hypothesis that students from the same college with equivalent UGPAs theoretically should obtain similar scores on the LSAT.[299]

Finally, we should expect that UGPAs with the same college to correlate with LSAT scores because law schools themselves make this assumption when formulating admissions policies. In an article titled "Toward a Sociology of Law School Admissions," Professor Howard Erlanger recommends that since "no organization follows all its own rules, the next question is how the policy is implemented."[300] Thus, the claim that the LSAT does not necessarily correspond with UGPA within the same college can be refuted by studying what law schools actually *do* in the admissions process, as opposed to what admission officials say. The evidence indicates that law schools and LSAC have acted upon the assumption of a positive, stable relationship between the LSAT and UGPA within a college in the manner they have dealt with the challenge of comparing the UGPAs of applicants from different institutions.

First LSAC, and then individual law schools, devised techniques for "adjusting" or "standardizing" grades from different colleges by using LSAT College Mean (LCM) data as a metric to calculate the comparative level of difficulty level in earning grades within an institution.[301] This issue

---

Jay Rosner, *supra* note 237, at app. (reporting unpublished study by David S. Mann and Thomas S. Gibson, who found that the LSAT correlated .82 with the SAT for College of Charleston students). The proportion of total variance accounted for by the association of two variables is calculated by squaring the correlation coefficient. Thus, a LSAT/FYA correlation of .40 accounts for 16% of variance, whereas a LSAT/SAT correlation of .72 accounts for 52% of variance, and a .86 correlation between the SAT and the GRE would account for 74% of variance. Even if one were to accept at face value LSAC's estimated correlation for range restriction, the LSAT/FYA correlation would still be .52, or 27% of variance. ANTHONY ET AL., *supra* note 155, at 10.

299.   *See* Gannon, *supra* note 5, at 274.

The assumption of little or no LSAT score difference between groups under the restricted conditions just outlined appears to be justified by the absence of any pure test of ability to perform in law school. Current attempts to conceptually distinguish aptitude from achievement are difficult enough; the practical development of distinguishable measures of each psychological construct is virtually impossible. Hence, performance on aptitude tests like the LSAT should closely approximate past scholastic achievement (indicated by grades) since aptitude tests for graduate and professional schools are intended to measure in large part the knowledge and intellectual skills acquired in school and college; and these are the factors that best predict success in specialized graduate programs.

Id.

300.   Howard S. Erlanger, *Toward a Sociology of Law School Admissions*, 34 J. LEGAL EDUC. 374, 383 (1984).

301.   *See* NAIRN, *supra* note 298, at 246-51 (describing how LSAC's Law School Data Assembly Service adjusted applicants UGPAs based on LCM data for about fifty law schools that requested the service, and how applicants were never informed, either by LSAC or individual law schools, that their grades were being altered); Guinier, *supra* note 222, at 579 n.43 (describing how LCM data was included in a mathematical formula used to rank and select applicants at the University of Pennsylvania School of Law); David Kaye, *An "A" is an "A" is an "A": An Exploratory Analysis of a New Method for Adjusting Undergraduate Grades for Law School Admissions Purposes*, 31 J. LEGAL EDUC. 233, 236-37 (1981) (describing a validation study of grade adjustments at the Arizona State University

was recently raised in the Fifth Circuit's December 2000 *Hopwood III* decision.[302] In its effort to show that Cheryl Hopwood and other plaintiffs in a "reverse discrimination" lawsuit would not have been admitted even under "race-blind" procedures, the University of Texas Law School relied on Professor Olin Wellborn of its faculty as an expert witness.[303] In his report, Wellborn acknowledged that file reviewers typically take account of the mean LSAT scores for the applicant's college, and that "comparing the colleges' mean LSAT scores is the best way of evaluating the various colleges and universities in terms of the caliber of their respective student bodies."[304] Thus, UGPA adjustments are anchored to LCMs; law schools and LSAC take it as axiomatic that an average LSAT score within a college can be expected to have a reliable equivalency with an average UGPA within that same college.[305]

In other words, though two individual students with the same UGPA in the same college may have appreciably different scores on the LSAT, grade adjustment techniques assume that across large samples of students, those with equivalent UGPAs in the same college should, on average, perform equally well on the LSAT. This assumption is similar to that operationalized in the present study. One difference is that rather than using LCM data as an indirect means of controlling for institutional differences, I directly controlled for institution by only matching applicants within the same college or university (and the same graduation dates and majors). In summary, none of the three alternative hypotheses evaluated in this section

---

School of Law); Kidder, *Rise of the Testocracy, supra* note 47, at 178-79 (describing the negative consequences for women of UGPA adjustments at UC Berkeley's Boalt Hall School of Law, where they were discontinued in 1998 amidst political controversy); Sander, *supra* note 80, at 492 n.39 (describing UGPA adjustments at the UCLA School of Law).

For LSAC sponsored national studies of the effect of UGPA adjustments on the prediction of first-year law school grades, which generally find negligible validity gains, see Robert F. Boldt, *Efficacy of Undergraduate Grade Adjustment for Improving the Prediction of Law School Grades, in* 3 REPORTS OF LSAC SPONSORED RESEARCH: 1975-1977, *supra* note 155, at 369; Robert F. Boldt & Russell A. Simpson, *Aptitude Moderation of Undergraduate Grades to Predict Law School Grades*, LSAC 78-3, *in* 4 REPORTS OF LSAC SPONSORED RESEARCH: 1978-1983, *supra* note 161, at 113; Donald A. Rock & Franklin R. Evans, *The Effectiveness of Several Grade Adjustment Methods for Predicting Law School Performance, in* 4 REPORTS OF LSAC SPONSORED RESEARCH: 1978-1983, *supra* note 161, at 363; W. B. Schrader & Barbara Pitcher, *Effect of Differences in College Grading Standards on the Prediction of Law School Grades, in* 3 REPORTS OF LSAC SPONSORED RESEARCH: 1970-1974, *supra* note 155, at 433.

302.   Hopwood v. Texas, 236 F.3d 256 (5th Cir. 2000). *Hopwood III* addressed various appeals and cross-appeals of the district court's rulings pursuant to remand in the historic 1996 *Hopwood II* Fifth Circuit decision, Hopwood v. Texas, 78 F.3d 932 (5th. Cir. 1996).

303.   *Hopwood*, 236 F.3d at 262.

304.   *Id.* at 268.

305.   Please note that this does not mean that I am arguing that UGPA adjustments necessarily make for wise admission policy. They can create artificial barriers to entry. Kidder, *Rise of the Testocracy, supra* note 47, at 174-79. The point here is merely to render certain logical consequences more visible by analyzing institutional practices.

seriously challenge the claim that either the LSAT or the testing milieu are culturally biased.

CONCLUSION

Many people, especially affirmative action critics and testing advocates, assume that standardized tests like the LSAT are a neutral reflection of racial and ethnic differences in educational achievement. This study of elite law school applicants, matched on UGPA within the same institutions and majors, establishes that such an assumption lacks empirical support.

Organizations active in the effort to dismantle affirmative action, such as the Center for Individual Rights (CIR), have developed their litigation strategies around the assumption that the LSAT equals merit.[306] Given the centrality of the testing issue to the affirmative action debate, it is essential that the fairness of the LSAT and other standardized tests be vigorously contested when "reverse discrimination" challenges to law school admission policies are still in the pretrial stage.[307] Unfortunately, universities'

---

306. *See* Expert Report of Kinley Lamtz, Grutter v. Bollinger, 16 F. Supp. 2d 797 (E.D. Mich. 1998) (No. 97-75928). Lamtz, an applied statistician who is serving as CIR's main expert in *Grutter*, attempted to establish the presence of impermissible racial preferences by calculating the odds of acceptance by racial group among applicants in the same small LSAT/UGPA bands. *Id.*; LAW SCH. ADMISSION COUNCIL, *supra* note 41, at 10.

> Opponents of affirmative action tend to view test scores as the sole or predominant measure to be used in the admission process. They have misused the LSAT most egregiously because their litigation theory clearly assumes that one who achieved a higher score on the LSAT is presumptively entitled to admission over one who receives a lower score.

Id. Lamtz's effort to substantiate that students of color at the University of Michigan Law School are "less qualified" than rejected White applicants is questionable. Professors Lempert, Chambers, and Adams argue:

> Although no one disputes the fact that LSAT scores and UGPAs are important in the selection of both white and minority law students, analyses like those of the plaintiffs in the Michigan lawsuit do not necessarily show that this is the case. Even when the odds that a minority student with certain hard credentials will be admitted are hundreds of times what they are for a white student with similar credentials, the evidence does not *necessarily* mean race was the crucial factor, or even that it figured in the admissions decision. As to any given student, it may be the case that some factor other than race, such as leadership ability, was the crucial factor in the decision.
>
> A thought experiment can make this clear. Suppose, for example, that neither minority status nor an LSAT/UGPA index counted in Michigan's admissions process, but students were instead admitted based on a factor, say proven leadership potential, which was similarly distributed among white and minority applicants and orthogonal to or only weakly correlated with the LSAT/UGPA index. Because most minorities in Michigan's applicant pool have index scores that are below the index scores of most whites in Michigan's applicant pool, admitting students with no attention to ethnicity could be expected to yield groups of minority and white matriculants whose index scores, on average, would differ substantially. . . . [L]ooking historically, differences in index scores do not necessarily tell us much about the degree to which ethnicity as opposed to letters of recommendation, a history of overcoming adversity, a history of outperforming standardized tests, leadership ability, impressions in a personal interview, or other factors dominated Michigan's admissions process or the admissions process in any law school.

Lempert et al., *supra* note 215, at 491-500 n.62.

307. *See* White, *supra* note 178. White argues:

institutional interests in student diversity only partly overlap with the interests of students of color in preserving educational access through affirmative action. Thus, in major affirmative action cases like *Bakke, DeFunis*, and *Hopwood*, universities' records and briefs were devoid of evidence that standardized tests like the LSAT are culturally biased against students of color.[308]

Grutter v. Bollinger has the potential for a more positive outcome. First, the University of Michigan is presenting scientific evidence that standardized tests unfairly penalize students of color.[309] Second, student of color and pro-affirmative action organizations, who, in contrast to *Hopwood*, have been permitted to intervene as defendants, are also mounting a spirited challenge to the fairness of the LSAT.[310]

Efforts to establish bias on standardized tests like the LSAT and SAT are particularly timely in the present litigation environment since the diversity rationale for affirmative action may soon be rejected or curtailed by the Supreme Court.[311] Supporters of affirmative action would be wise to cover

---

The importance of the issues raised by invalid admissions criteria and culturally biased testing demands that litigators and courts collect all the available evidence and give their fullest consideration to the facts revealed by such evidence. An effort commensurate with that in *Brown v. Board of Education* should be made to bring the country's finest minds, from all the relevant sciences, to bear on the problem. And this effort should occur as a preparation for trial, not appeal, of the issues: Evidence of cultural bias must be adduced *on the record* to persuade appellate courts fully to consider the issues and, hopefully, to resolve them in favor of recognizing and correcting for predictive invalidity and cultural bias in standardized tests.

Id. (footnote omitted).

308.   *See id.* at 87, 124-25 (arguing that the universities party to *Bakke* and *DeFunis* declined to present evidence of bias in testing for fear that this would expose them to litigation from rejected minority applicants); JOEL DREYFUSS & CHARLES LAWRENCE III, THE BAKKE CASE: THE POLITICS OF INEQUALITY 32, 39-53 (1979) (criticizing the UC's restricted defense of affirmative action, including how UC had a greater institutional interest in quickly establishing a clear precedent rather than in mounting a defense under optimal conditions, how UC neglected to argue that Bakke's admission may have been effected by the Davis dean's discretionary admissions prerogative for the wealthy and connected, and how UC contributed to the notion that Bakke was better qualified based on test scores and undergraduate grades); Brown-Nagin, *supra* note 85, at 380-81, 393-94 (criticizing the *Hopwood* decision, and noting a failed effort by the NAACP Legal Defense Fund and the Thurgood Marshall Law Student Association to intervene on grounds that the UT would not adequately litigate issues of bias on the LSAT out of concern over Title VI liability).

309.   Claude M. Steele is an expert witness in both the *Gratz* and *Grutter* cases. For analysis of Steele's research on stereotype threat, see *supra* Part IV.B. Nevertheless, legal counsel for the University of Michigan Law School decided not to have Steele testify in the *Grutter* trial even though bias in admission criteria was one of the issues Judge Bernard Friedman wanted to address in the trial.

310.   See *supra* Part IV.C.4 for discussion of the intervenors' LSAT expert witnesses in the *Grutter* case.

311.   *See* Wessmann v. Gittens, 160 F.3d 790, 797-98 (1st Cir. 1998) (holding that even if *Bakke* is still good law, that the diversity rationale cannot justify race-based affirmative action beyond a small threshold); Hopwood v. Texas 78 F.3d 932, 944 (5th Cir. 1996).

We agree with the plaintiffs that any consideration of race or ethnicity by the law school for the purpose of achieving a diverse student body is not a compelling interest under the Fourteenth Amendment. Justice Powell's argument in *Bakke* garnered only his own vote and has never represented the view of a majority of the Court in *Bakke* or any other case.

their bases by supplementing their defense of the diversity rationale with proof of prior discrimination[312] or evidence of bias in standardized tests and other traditional admission criteria.

If race-conscious affirmative action is declared constitutionally impermissible, the debate over the fairness of the LSAT and other standardized tests becomes more, not less, important. Professor Daria Roithmayr, for example, has identified a formidable network of anticompetitive institutional relationships that has locked in standards of merit governing entrance to law school that favor Whites.[313] Without the momentum of an energetic appraisal of traditional merit criteria like the LSAT, there is little hope of altering this maze of feedback loops, which range from the *US News & World Report* rankings to ABA accreditation requirements, in order to promote more equitable opportunities to enter the legal profession.

Concerns over ethnic bias in standardized tests need to be linked to a more far-reaching and transformative critique of the conventional higher education testocracy.[314] For instance, Professor Lani Guinier, in reviewing

---

Id.; Tracy v. Bd. of Regents of the Univ. Sys. of Ga., 59 F. Supp. 2d 1314, 1321-23 (S.D. Ga. 1999) (implying that the diversity rationale may not be a compelling governmental interest); Delgado & Stefancic, *supra* note 8, at 1528-29 (reviewing recent setbacks to Justice Powell's diversity rationale in *Bakke*); *Race-Sensitive Admissions in Higher Education: Commentary on How the Supreme Court Is Likely to Rule*, 26 J. BLACKS HIGHER EDUC. 97 (1999-2000) (interviewing five legal scholars, all of whom separately predicted that the Court would restrict or reject the diversity rationale from *Bakke*).

312.   *See* Delgado & Stefancic, *supra* note 8, at *passim* (reviewing discrimination against minorities in California); Richard Delgado & Jean Stefancic, *Home-Grown Racism: Colorado's Historic Embrace—and Denial—of Equal Opportunity in Higher Education*, 70 U. COLO. L. REV. 703 (1999) (reviewing discrimination against minorities in Colorado).

313.   Daria Roithmayr, *Barriers to Entry: A Market Lock-in Model of Discrimination*, 86 VA. L. REV. 727 (2000). Roithmayr argues:

Institutional networks of interdependent professional organizations have now grown up around this culturally specific standard to lock the standard in. To gain access to the legal profession network, for example, law schools have had to adopt the industry standard that favors whites. For example, schools that want to maintain their national ranking or place graduates in lucrative positions must admit students based on their Law School Admission Test ("LSAT") scores, because legal professionals recognize the test as a signal of "quality." Over time, the standard has become progressively more embedded into the network, creating significant barriers to entry for people of color whose cultural performances do not conform to the standard.

Id. at 34; *see also* Kidder, *Portia Denied, supra* note 47, at 22-23 (detailing the misuse and overreliance on LSAT scores in law school admission decisions, financial aid allocation, selection of prestigious federal clerkships, attorney hiring decisions, and university funding).

314.   *See* Sturm & Guinier, *supra* note 78, at 956. Sturm and Guinier contend:

It is time, we argue, for those of us committed to racial and gender equity to advance a more fundamental critique of existing selection and admission conventions. It is time to discuss how conventional assessment and predictive criteria do not function fairly, democratically, or even meritocratically for many Americans who are not members of racial or gender minorities. To reclaim the moral high ground, we must broaden and expand the terms of engagement. By revealing faulty assumptions about the concept of affirmative action and the system of selection in which it operates, we can move from an incrementalist strategy of inclusion for a few to a transformative vision of reform for the many.

Id.

the recent comprehensive study of 1970-96 Michigan Law School alumni by Lempert, Chambers, and Adams,[315] concludes:

> The study confirms the benefits of affirmative action to all Michigan graduates. It tells us that affirmative action critics' much-touted reliance on *objective measures of merit* have little to recommend them over the life span of a lawyer. After all, it is the life's work of the graduates that is the big test. Thus, rather than ban affirmative action, the law school might do well to expand its practice and to revamp the admissions criteria for *all* incoming law students.[316]

The Michigan study, to the degree that it can be generalized to other peer institutions, supports the position that law school admission criteria should be revamped. The authors found no relationship between LSAT/UGPA index scores and subsequent success in the legal profession, as measured by income or career satisfaction.[317] Moreover, the minority graduates of Michigan, who had appreciably lower average LSAT scores, nonetheless went on to serve as leaders in public service at higher rates than their White classmates.[318] The hidden societal costs of selection criteria heavily dependent on the LSAT extend beyond the Michigan study. Research shows a negative correlation between social activism and performance on the LSAT for the national pool of test takers.[319]

Given the results presented in this study, it should not come as a surprise that in the four years since Proposition 209 and SP-1, 25% of White applicants were accepted to Boalt Hall, UCLA Law School and UC Davis Law School, compared to 21% of Asian Americans, 18% of Chicanos and Latinos, 15% of Native Americans and only 11% of African Americans.[320] Yet in California's stark post-affirmative action landscape there is a possibility worth exploring. It is always important to recognize the myriad historical and institutional barriers that have inhibited the educational attainment of students of color in the United States. However, the data reported in this study suggest that the LSAT decisively stratifies opportunity by race even among law school applicants who have (sometimes

---

315. Lempert et al., *supra* note 215, at 395.
316. Guinier, *supra* note 222, at 582.
317. Lempert et al., *supra* note 215, at 468.
318. *Id.*, at 485-89.
319. Kidder, *Rise of the Testocracy*, *supra* note 47, at 202 (reviewing national studies conducted by ETS and UCLA); *see also* Guinier, *supra* note 222, at 574.

> One can certainly begin to speculate, however, that multiple-choice, timed testing may train successful candidates not to question authority, not to look for innovative ways to solve problems, not to do sustained research or to engage in team efforts at brainstorming, but instead to try to answer questions quickly and in ways that anticipate the desires or predilections of those asking the questions.

Id.

320. UC Office of the President, *supra* note 71. These figures are separately reported by school, so applicants to more than one UC law school are duplicated. *Id.* at n.1.

poignantly) overcome obstacles to achieve equivalent academic success over four or more years of college. Given that the LSAT extracts such a substantial "pioneer tax" on many of the most accomplished minority law school candidates in the country, it might be argued that the UC law schools violate Proposition 209's prohibition against racial and ethnic preferences[321] by relying so heavily on the LSAT.[322]

One final point concerns Professors Jencks and Phillips's suggestion, mentioned at the beginning of this Comment, that reducing the Black-White test score gap would do more to promote racial equality than any other feasible strategy.[323] There may indeed be less racial inequality if performance differences on standardized tests suddenly shrank, but it does not follow that America's best chance to decrease racial inequality lies in closing the test score gap. Since, at least in the law school context, the Black-White and Latino-White gaps are greater on the LSAT than on undergraduate grades or law school grades, the continued emphasis on the LSAT acts as an artificial barrier for students of color aspiring to enter the legal profession. Thus, organizing social policy around test scores on the false assumption that tests like the LSAT are a neutral measure of educational differences can exacerbate, rather than lessen, racial and ethnic inequality. Jencks and Phillips's observation that it will take "several generations" for African Americans to adjust to "spending more time studying"[324] implicitly amounts to an "all deliberate speed"[325] approach to ending racial stratification. Even worse, it adds insult to injury to the large number of students of color who have already matched the performance of

---

321.   *See* Richard Delgado, Official Elitism or Institutional Self-Interest? 10 Reasons Why UC Davis Should Abandon the LSAT (and Why Other Good Law Schools Should Follow Suit), Barrett Lectureship on Constitutional Law at UC Davis Law School (Oct. 12, 2000) (arguing that heavy reliance on the LSAT may violate Proposition 209). This lecture will be reprinted in 34 U.C. DAVIS L. REV. (forthcoming 2001).

322.   Kidder, *Rise of the Testocracy, supra* note 47, at 193 (reviewing the role of the LSAT in admissions to the law schools at UC Berkeley, UCLA, Davis, and Hastings); Cecilia Estolano et al., New Directions in Diversity: Charting Law School Admissions in a Post-Affirmative Action Era 26 (1997) (unpublished manuscript, on file with the Boalt Hall Law Library) (reporting that a one-point difference on the LSAT made the difference between a thirteen percent chance of admission and a thirty-one percent chance of admission to Boalt Hall).

323.   Jencks & Phillips, *supra* note 21, at 4-5.

324.   Christopher Jencks & Meredith Phillips, *Black-White Test Score Gap Is Not Inevitable*, L.A. TIMES, Sept. 28, 1998, at B5; *see also* Christopher Jencks & Meredith Phillips, *America's Next Achievement Test: Closing the Black-White Test Score Gap*, 40 AM. PROSPECT (1998), *at* http://prospect.org/print/V9/40/jencks-c.html (arguing that African American culture must adjust to the premium placed on cognitive skills in the new economy).

> [I]t always takes several generations for any group to adjust to a new reality, especially when the adjustment has significant costs (spending more time studying). The message that nerds will do well as adults is always hard to sell to children, but it is doubly hard to sell when it has only recently become true.

Id.

325.   Brown v. Bd. of Educ., 349 U.S. 294, 301 (1955).

their White classmates in college, and yet are suddenly and systematically made to appear "less qualified" after taking a four-hour LSAT.