Michigan Journal of Race and Law

Volume 11

2006

The Diversity Rationale: Unprovable, Uncompelling

Brian N. Lizotte
Yale Law School

THE DIVERSITY RATIONALE: UNPROVABLE, UNCOMPELLING

*Brian N. Lizotte**

Student body diversity—and the purported educational benefits diversity bestows—is the final Supreme Court-endorsed justification for affirmative action by public universities. Are the benefits of diversity indeed "substantial," as the Grutter *majority claimed? The author analyzes the social scientific research upon which the Court relied in articulating the diversity interest. By critiquing its theory and methodology, the author shows how the research fails to prove educational benefits; and by considering the logic underlying social science generally, he shows how the causal relationship is, technically, not provable. The author questions, then, how the diversity interest can possibly be compelling.*

---

* B.A. & M.S., Yale College; Ed.M., Boston University; J.D., Yale Law School. Thank you to Stephen Yandle and David Peters for commenting; to Alan Schoenfeld for assisting my research; to my lead editor, Sacha Montas; to Peter Salovey, Brian and Jerusha Detweiler-Bedell for teaching me; and, of course, to my wife, Lee, and my family.

INTRODUCTION

Student body diversity is now endorsed by a Supreme Court majority as a compelling state interest—generating unique educational benefits—for which a public university may strive by means of a narrowly tailored affirmative action program.[1] *Grutter* and *Gratz* affirm Justice Powell's solitary but determinative acceptance in *Bakke* of the use of race by the University of California to attain a diverse student body.[2] Indeed, the supposed educational benefit of diversity is now the *only* permissible basis for racial preferences in college admissions; *Bakke* became a "shibboleth,"[3] foreclosing corrective and equal access justifications.[4]

In *Bakke*, Powell recognized a university's right to "select those students who will contribute the most to the 'robust exchange of ideas,'" a goal of "paramount importance in the fulfillment of [the university's] mission."[5] A diverse enrollment was intended to create an atmosphere "'most

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

2. Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265, 311 (1978). Several appeals courts rejected Powell's lone-authored opinion as binding precedent. Johnson v. Bd. of Regents of the Univ. of Ga., 263 F.3d 1234, 1261 (11th Cir. 2001) ("[W]e do not believe that Justice Powell's opinion is binding[.]"); Hopwood v. Tex., 78 F.3d 932, 944 (5th Cir. 1996) ("Justice Powell's view in *Bakke* is not binding precedent on this issue. While he announced the judgment, no other Justice joined in that part of the opinion discussing the diversity rationale.").

3. John H. Bunzel, *The Diversity Dialogues in Higher Education*, 29 FORDHAM URB. L.J. 489, 500 (2001).

4. In his separate *Bakke* opinion, Justice Marshall warned of the chilling effect that Powell's foreclosure would entail. *Bakke*, 438 U.S. at 399 (Marshall, J., separate opinion) ("As we have observed, '[a]ny other approach [than remediation] would freeze the status quo that is the very target of all desegregation processes.'") (quoting McDaniel v. Barresi, 402 U.S. 40, 41 (1971)). *See* Charles R. Lawrence III, *Two Views of the River: A Critique of the Liberal Defense of Affirmative Action*, 101 COLUM. L. REV. 928, 931 (2003) ("[A]s diversity has emerged as the dominant defense of affirmative action in the university setting, it has pushed other, more radical substantive defenses to the background.").

5. *Bakke*, 438 U.S. at 313 (quoting U.S. v. Associated Press, 52 F. Supp. 362, 372 (S.D.N.Y. 1943)).

conducive to speculation, experiment, and creation,' "[6] preparing students, in turn, to live in a diverse society. Powell claimed, "[O]ur tradition and experience lend support to the view that the contribution of diversity is substantial."[7] Powell confined his argument to "tradition and experience," offering no formal data to support his expostulation.

In the years between *Bakke* and *Grutter/Gratz*, the Court clarified its affirmative action jurisprudence in two ways. First, the Court, through three opinions written by Justice O'Connor, crystallized a single, strict scrutiny standard for all racial classifications, even those having a supposedly benign purpose. Following her concurrence in *Wygant* (sharing in "the belief, apparently held by all Members of this Court, that racial classifications of any sort must be subjected to 'strict scrutiny,' however defined"),[8] three years later, O'Connor expanded her defense of a singular strict scrutiny standard in the *Croson* majority opinion;[9] and finally, in 1995's *Adarand Constructors*, she answered firmly any doubt lingering post-*Croson* by asserting, "[W]e hold today that all racial classifications, imposed by whatever federal, state, or local governmental actor, must be analyzed by a reviewing court under strict scrutiny."[10]

Having firmly established a strict scrutiny standard, the Court next winnowed away the remedial justification for affirmative action. Justice Brennan had originally argued in his separate *Bakke* opinion that affirmative action was an appropriate policy response to correct for the broad history of discrimination in American society.[11] However, that view was never shared by a Court majority, and in *Croson*, Justice O'Connor explicitly rejected the broad societal theory, holding that affirmative action was permissible only as a remedy for specifically identified, purposeful discrimination by a particular actor, and only for those particular groups who had experienced the discrimination.[12] Thus, a school that could not, or was unwilling to, show that it had previously discriminated against particular racial groups was now left with only one justification for employing racial preferences in its admissions decisions: the educational

---

6. *Id.* at 312 (quoting Sweezy v. N.H., 354 U.S. 234, 263 (1957) (Frankfurter, J., concurring in result)).

7. *Id.* at 313.

8. Wygant v. Jackson Bd. of Educ., 476 U.S. 267, 285–86 (1986) (O'Connor, J., concurring in part and concurring in the judgment).

9. City of Richmond v. J.A. Croson Co., 488 U.S. 469 (1989).

10. Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 227 (1995).

11. *Bakke*, 438 U.S. at 325 (Brennan, J., concurring in the judgment in part and dissenting in part) ("Government may take race into account when it acts not to demean or insult any racial group, but to remedy disadvantages cast on minorities by past racial prejudice."); *id.* at 362 ("[The University's] articulated purpose of remedying the effects of past societal discrimination is, under our cases, sufficiently important to justify the use of race-conscious admissions programs where there is a sound basis for concluding that minority underrepresentation is substantial and chronic[.]").

12. *Croson*, 488 U.S. at 493.

benefits of diversity. But Justice O'Connor rejected even this "too amorphous, too insubstantial" interest,[13] at least until 2003.

Grutter and *Gratz* finally took on the diversity rationale directly, striving to imprint *Bakke*'s hunch about the educational benefits of diversity with actual social scientific evidence. Although deferring to the University of Michigan Law School's "educational judgment that such diversity is essential to its educational mission"[14]—a generous deferral—the *Grutter* Court concluded in its own right, based on evidence provided by *amici curiae*, that "the educational benefits that diversity is designed to produce ... are substantial."[15] So definitive a statement proclaims the diversity rationale as fact, certifying Justice Powell's tempered words.[16] Justice O'Connor's reliance on social science departs from her past warning that "[s]ocial scientists may debate how peoples' thoughts and behavior reflect their background, but the Constitution provides that the Government may not allocate benefits and burdens among individuals based on the assumption that race or ethnicity determines how they act or think."[17] Nonetheless, in *Grutter*, she finds that diversity, by itself 1) creates cross-racial understanding; 2) breaks down racial stereotypes; 3) generates classroom discussion that is " 'livelier, more spirited, and simply more enlightening and interesting' "; and 4) prepares students for the workforce and society.[18]

---

13. Metro Broad., Inc. v. F.C.C., 497 U.S. 547, 612 (1990) (O'Connor, J., dissenting). *Accord Croson*, 488 U.S. at 493 ("Classifications based on race carry a danger of stigmatic harm. Unless they are *strictly reserved* for remedial settings, they may in fact promote notions of racial inferiority and lead to a politics of racial hostility.") (emphasis added) (citing *Bakke*, 438 U.S. at 298).

14. *Grutter*, 539 U.S. at 328. This deference is grounded in a university's academic freedom, protected by the First Amendment. *Bakke*, 438 U.S. at 312 ("Academic freedom, though not a specifically enumerated constitutional right, long has been viewed as a special concern of the First Amendment. The freedom of a university to make its own judgments as to education includes the selection of its student body."). Could not "academic freedom" also be used to justify a university's decision to admit only White students, in service of sincerely-held ideological and pedagogical beliefs concerning White supremacy? Charles R. Lawrence III, *Each Other's Harvest: Diversity's Deeper Meaning*, 31 U.S.F. L. REV. 757, 770–71 & nn.70–74 (1997). Seemingly not. *Cf.* Bob Jones Univ. v. United States, 461 U.S. 574 (1983) (holding that the IRS may deny tax-exempt status to private schools who practice racial discrimination, but who would otherwise qualify for an exemption. Bob Jones University is dedicated to teaching fundamentalist Christian religious beliefs, and was, consistent with these beliefs, discriminating against non-White students in admissions. The Court held that there was an overriding government interest in eradicating racial discrimination that outweighed the university's free exercise of religion). "Academic freedom" is apparently handcuffed when it resists rather than promotes racial diversity.

15. *Grutter*, 539 U.S. at 330. Twelve expert witnesses in *Gratz* and fourteen in *Grutter* testified on the costs and benefits of considering race in admissions. Several *amici* briefs are available at http://www.umich.edu/~urel/admissions/legal (last visited Mar. 7, 2006).

16. *See supra*, text accompanying note 7.

17. *Metro Broad., Inc.*, 497 U.S. at 602 (O'Connor, J., dissenting).

18. *Grutter*, 539 U.S. at 330 (quoting App. to Pet. for Cert. 244a, 246a).

This Note confronts the social scientific research upon which the *Grutter* Court relied in articulating that a diverse student body yields educational benefits. By criticizing its theory and methodology, I show how the research fails to prove educational benefits; and by considering the logic underlying social science generally, I show how the causal relationship is, technically, not provable. Social science can only speak to what *might* be true. Ultimately, then, the diversity interest is an unstable foundation for affirmative action policy—certainly not compelling.

I. SOCIAL SCIENCE: LOGIC AND ROLE IN EQUAL PROTECTION LAW

Social science is grounded in probability, not deductive logic. Social scientists provide evidence of an effect only by ruling out the possibility that there is *no* effect. Consider a hypothetical researcher who finds that Black students at colleges with minority populations exceeding 40% score, on average, 155 points on the LSAT, while Black students at colleges with minority populations below 40% score only 148 points. She reports with "95% confidence" that being educated alongside a diverse student body improves Black students' LSAT performance. This conclusion is inaccurate in two important ways.

First, complementary to her 95% confidence is a 5% chance of error. Social scientists do not prove the truth of a proposition; instead, they indirectly show that a proposition is *likely* by showing that a competing "null" hypothesis is *unlikely*. In the example, the researcher would ask, "What is the chance, given the null hypothesis that student body diversity does *not* affect Black students' LSAT scores, that, in my particular sample, I would actually find a seven point advantage for Black students educated alongside a more diverse student body?" The question is cumbersome, but the answer is simple: 5%. That is, the researcher reasons that it is sufficiently *un*likely (only 5%) that she would find a seven point difference in her sample if the true population difference were zero,[19] so she *infers* (but does not prove) that campus diversity has a genuine effect on Black students' LSAT scores.

This is the best she can do; she can never be certain. A critic's job is comparatively easy,[20] because there is always some chance the seven-point difference was due to random chance, unrelated to diversity. Given the

19. A "population" is the set of *all* people possessing a given characteristic; here, all Black college students worldwide who take the LSAT. A "sample" is the subset of the population that a researcher actually studies. From conclusions reached by studying her sample directly, a researcher makes inferences about the population in general.

20. *E.g.*, Crystal G. Muhammad, Data Matters: Making a Compelling Case for Diversity in Education 22 (2003) (unpublished Ph.D. dissertation, University of Virginia) (on file as UMI Microform 3091141 with ProQuest Information and Learning Company, Ann Arbor, MI) (complaining that social scientific studies are difficult to conduct but easy to criticize).

impossibility of proof by social science, the *Grutter* majority's proclamation that "the educational benefits that diversity is designed to produce ... are substantial"[21] is not phrased with requisite caution. The Court proclaimed a compelling interest in the benefits of student body diversity only by relying on evidence that it is unlikely diversity has *no* effect. Worse, the evidence came from methodologically suspect research.

A second problem for my hypothetical researcher is that, even if she could be 100% certain of a genuine difference in Blacks' LSAT scores between more and less diverse schools, such a finding would explain nothing about how, why, or even *if* diversity caused the difference. Studies that take student samples as they find them—neither randomly selecting students from a larger population nor randomly assigning them to treatment conditions—can only detect correlations, not causation. It is always possible to "explain away" a correlation. For instance, schools with large endowments may be able to afford both greater diversity through generous financial assistance, and richer educational opportunities through opulent resources. Then, the seven point advantage in my hypothetical might be a pure resource effect,[22] with no role played by campus diversity. To infer that diversity *causes* an educational benefit, one would need to randomly select hundreds or thousands of Black students, randomly assign one half to an all-Black school and the other half to a racially heterogeneous school, and extensively control other extraneous variables on each campus for four years, in order to isolate a diversity effect. It is unsurprising that no such study has been performed, and unlikely one ever will be performed.

Numerous scholars have argued against using social science to decide equal protection claims.[23] First, they argue that courts, by relying on social science, demean the moral principle embodied in the Fourteenth Amendment.[24] Today, the *Grutter* Court argues, social science demonstrates a reliable effect of diversity on educational outcomes. What if, twenty-five years from now, the social science changes, and the purported educational benefit of diversity disappears? Would affirmative action suddenly be unconstitutional?

A related criticism is that it is institutionally inappropriate for judges to evaluate social science. Because social science is never flawless, judges act as legislators when they use such evidence to choose a particular side.

---

21. *Grutter*, 539 U.S. at 330.

22. *See infra* note 41.

23. *See generally* Scott Jaschik, *A Valuable Tool or Bias in Reverse?*, CHRON. OF HIGHER EDUC., Apr. 28, 1995, at A14.

24. Edmund Cahn, *Jurisprudence*, 30 N.Y.U. L. REV. 150, 167 (1955) ("It is one thing to use the current scientific findings ... in order to ascertain whether the legislature has acted reasonably" but "quite another thing to have our fundamental rights rise, fall or change along with the latest fashions of psychological literature."); Deborah J. Merritt, *The Future of* Bakke: *Will Social Science Matter?* 59 OHIO ST. L.J. 1055, 1056–57 (1998).

A federal court may decide whether Congress has acted reasonably in enacting a law, but a court traditionally does not weigh evidence itself where its decision will impact similarly situated litigants.[25] Many judges lack backgrounds in economics, sociology, and psychology, and may be unable to distinguish between methodologically sound and suspect research.[26] Judges interpreting social science may base their decisions not in preexisting legal principles, but in the judges' personal beliefs about the worthiness of a governmental objective.[27] Indeed, it is impossible to know whether a court citing social science data is relying on those data as necessary to its decision, or merely as supporting a decision reached on other grounds.

Two famous Supreme Court cases—*Brown v. Board of Education*[28] and *Muller v. Oregon*[29]—illustrate, in my opinion, the proper versus improper use of social science data to resolve an equal protection claim. In *Muller*, the Court for the first time relied on historical and anthropological data to justify the proposition that a woman's physical structure and role as mother "properly placed [her] in a class by herself," dependent on men, and creating a public health concern justifying limitation of her workday hours.[30] Today, of course, *Muller* reads as blatantly misogynistic, its "data" dated and disproved. Social science, by incorporating evolving cultural norms, is an unstable foundation for policy. The Court's difficulty came in trying to use social science for something it is ill-equipped to do: prove a proposition.

In contrast, future generations have embraced that most famous use of social science in a Supreme Court case, the evidence cited in *Brown*[31] to

---

25. Muhammad, *supra* note 20, at 29–30 (calling "most controversial" judges' "legislative use of social science to create new law . . . applicable to every similarly situated set of plaintiffs and defendants.").

26. J. HARVIE WILKINSON, III, FROM *BROWN* TO *BAKKE*, THE SUPREME COURT AND SCHOOL INTEGRATION: 1954–1978 32 (1979); Russell B. Korobkin & Thomas S. Ulen, *Law and Behavioral Science: Removing the Rationality Assumption From Law and Economics*, 88 CAL. L. REV. 1051, 1093–94 (2000).

27. Robert P. George, Gratz *and* Grutter*: Some Hard Questions*, 103 COLUM. L. REV. 1634, 1637 (2003) (judges might "vary in their judgments not because of differences of opinion about law, but purely as a result of ideological differences."); James E. Ryan, *The Limited Influence of Social Science Evidence in Modern Desegregation Cases*, 81 N.C. L. REV. 1659, 1676–80 (2003) (observing that judges are susceptible to confirmation bias, interpreting evidence in a manner that confirms their initial beliefs, and discounting evidence that contradicts their beliefs).

28. 347 U.S. 483 (1954).

29. 208 U.S. 412 (1908).

30. *Id.* at 422. The Court relied principally on the "Brandeis Brief," consisting of two pages of legal reasoning and 100 pages of social scientific research.

31. 347 U.S. at 494–95, n.11. Chief Justice Warren cited seven social scientific authorities, principally KENNETH B. CLARK, EFFECT OF PREJUDICE AND DISCRIMINATION ON PERSONALITY DEVELOPMENT (Midcentury White House Conference on Children and Youth, 1950). Muhammad, *supra* note 20, at 21–22, summarizes critiques of the Clark

attack *Plessy v. Ferguson*'s "separate but equal" doctrine justifying segregation.[32] The key difference between *Brown* and *Muller* is that the *Brown* Court used social science to *defeat* a policy ("separate but equal") by falsifying its assumptions.[33] The *Muller* Court, instead, erred by using social science to *justify* a policy (limitations on women's working hours) by establishing its assumptions. However, a single counterexample (e.g., a brawny, childless woman) negates the assumptions, causing the general policy to fail. Consistent with its logic, social science can demonstrate the falsehood, but never the truth, of a proposition.

By employing social science to articulate a diversity interest justifying affirmative action policies, the *Grutter* Court acted like the *Muller* Court, basing its decision necessarily on falsifiable evidence. I now turn to that evidence.

II. DIVERSITY RATIONALE: EVIDENCE BEFORE *GRUTTER*

Until the drive produced by the Michigan cases, there was little empirical research on educational outcomes related to affirmative action.[34]

---

study, including that the study used too small a sample, failed to control for Blacks' geographical migration patterns during the 1930s and 1940s, ignored the effects of question sequencing, and overstated its own results.

32. 163 U.S. 537 (1896). In *Plessy*, the Court upheld a Louisiana statute used to eject a U.S. citizen of 1/8 African and 7/8 Caucasian heritage from a first class railroad car. The Court protected the liberty of the Louisiana General Assembly to crystallize social norms, to "act with reference to the established usages, customs, and traditions of the people, and with a view to the promotion of their comfort." *Id.* at 550.

33. Several authors argue that moral principle, not social science, drove the *Brown* decision. RICHARD KLUGER, SIMPLE JUSTICE 706 (1975) ("Then [Warren] added, by way of stressing that the sociology was merely supportive and not the substance of the holding, 'It was only a note, after all.'"); Muhammad, *supra* note 20, at 4 ("For Warren, the moral principle of equality dictated desegregation, without regard for the judicial means to arrive at that end. . . . [T]he scholarship of legal historians as well as Warren's own footnote reference to the Clark evidence suggest that the research was not central to his reasoning."). James Ryan writes:

> [I]t is difficult to reconcile the notion that social science evidence was determinative in *Brown* with the fact that the Court relied on its decision in *Brown*, and nothing more, to outlaw segregated golf courses, buses, and beaches. . . . [A]ccounts of those who drafted the opinion, as well as an understanding of the political history surrounding the decision, strongly suggest that the evidence was cited to bolster and obfuscate what was at the time a fairly controversial normative conclusion that segregation . . . was morally wrong.

Ryan, *supra* note 27, at 1665–66 (citations omitted). *See generally* Sanjay Mody, Note, *Brown Footnote Eleven in Historical Context: Social Science and the Supreme Court's Quest for Legitimacy*, 54 STAN. L. REV. 793, 811–28 (2002)

34. William C. Kidder, *Affirmative Action in Higher Education: Recent Developments in Litigation, Admissions and Diversity Research*, 12 BERKELEY LA RAZA L.J. 173, 221 (2001).

Research was, and continues to be, politicized.[35] In this section, I review hallmarks in the research preceding *Grutter* and *Gratz*, including several studies whose authors served as expert witnesses in the *Grutter* and *Gratz* trials. The evidence lends itself to multiple interpretations. Such ambiguity reconciles poorly with a rigid compelling interest standard so strongly biased against any use of race in government programs.

A. *Evidence from Primary and Secondary Schools*

In 1998, Maureen Hallinan reviewed "well-designed and well-executed" studies on the educational benefits of diversity.[36] She does not elaborate on what it means to be well-designed and executed, except that the studies were "based on high quality data sets, rely on appropriate research designs, and employ rigorous analytic techniques,"[37] three vague criteria. Hallinan limits her focus to the effects of racial diversity. Even though *Grutter* urges attention to all types of diversity, it accepts the Law School's "longstanding commitment to 'one particular type of diversity,' that is, 'racial and ethnic diversity.' "[38] Thus, Hallinan's review is relevant.

Overall, Hallinan concludes, "[W]hile the research reveals a few inconsistencies, the major results receive wide empirical support" for diversity's benefit.[39] However, the evidence suffers from flaws and ambiguities, and can readily be explained by causes separate from diversity. Importantly, her review contains no randomized, controlled experiment testing the educational benefits of campus diversity, so it adds nothing to our knowledge of the causal relationship, if any exists, between the two variables.

Most of Hallinan's data come from elementary and secondary schools. She reports that Black, Latino, and White students all perform better academically in majority White schools than in minority White schools.[40] But this finding is hardly reducible to an effect of diversity, and is more easily explained in terms of systematic differences in resources available to majority White versus minority White schools, and the

---

Most inquiries into the educational benefits of diversity focused on elementary and secondary schools, making suspect any generalization to college populations, who are the more likely subjects of affirmative action policy.

35. *E.g.*, Ryan, *supra* note 27, at 1675 ("[A] number of social scientists studying desegregation seem precommitted to particular findings. One can often predict the conclusions of a report based on the identity of the author.").

36. Maureen T. Hallinan, *Diversity Effects on Student Outcomes: Social Science Evidence*, 59 OHIO ST. L.J. 733, 741 (1998).

37. *Id.* at 753.

38. *Grutter*, 539 U.S. at 316 (quoting the Law School's admissions policy). *But see infra* notes 102–108 and accompanying text.

39. Hallinan, *supra* note 36, at 741.

40. *Id.* at 741–42.

disparate educational opportunities these resources create.[41] Where a handful of Black students are performing well in a wealthy school alongside a thousand high-performing White students, it would be silly (and possibly insulting) to attribute the Black students' achievements to their diverse educational experience. Hallinan admits, "[I]t is not desegregation *per se* that improves achievement, but rather the learning advantages some desegregated schools provide."[42] Likewise, her observation that Black students' elevated achievement corresponds to how early they are placed in majority White schools[43] is more plausibly explained by longer exposure to a lion's share of resources than by the effects of integration.

Next, Hallinan observes that cooperative learning techniques are shown to increase achievement of all students in racially heterogeneous groups.[44] Here, it is impossible to separate the effect of a racially heterogeneous group from the effect of cooperative learning endeavors generally. Cooperative learning may be a successful technique independent of who is cooperating, and an established cooperative learning program should be equally successful in a classroom of all White or all Black students.

The most coherent of Hallinan's findings is that "research is fairly consistent in reporting that black and white students in desegregated schools are less racially prejudiced than those in segregated schools.... [I]nterracial contact in desegregated schools leads to an increase in interracial sociability and friendship."[45] The relationship is intuitive, and, although not fool-proof, the causal inference between exposure and friendship seems irrefutable. Whether or not interracial sociability is a bona fide "educational" benefit is less clear.

B. *Evidence from Universities*

Studies of the relationship between college student body diversity and educational benefits are comparatively sparse. Hallinan reports one

---

41. For a recent comprehensive analysis of the high correlation between academic achievement and family socio-economic status, *see* SUSAN E. MAYER, WHAT MONEY CAN'T BUY: FAMILY INCOME AND CHILDREN'S LIFE CHANCES (1997). *See also* David M. Engstrom, *Civil Rights Paradox? Lawyers and Educational Equity*, 10 J.L. & POL'Y 387, 412 (2002) (citations omitted) (explaining that income and education create "social capital," in turn improving educational outcomes). State funding inequities are tied to race and socioeconomic status. Tom Owens & Jeffery Maiden, *A Comparison of Interschool and Interdistrict Funding Equity in Florida*, 24 J. OF EDUC. FIN. 503 (1999). Teachers flock from impoverished urban schools with low concentrations of White students to wealthier suburban districts with high concentrations of White students. Leanna Stiefel et al., *Intra-District Equity in Four Large Cities: Data, Methods, and Results*, 23 J. OF EDUC. FIN. 447 (1998).

42. Hallinan, *supra* note 36, at 744.

43. *Id.*

44. *Id.*

45. *Id.* at 745.

study[46] showing an advance in "cognitive development" among students enrolled in a college course on multiculturalism, and a second study[47] showing that White students enrolled in a college diversity course "increase their understanding of the concept of race." Unfortunately, both studies were presented at a national conference and are unpublished; Hallinan describes them no further. On the surface, the findings are meaningless. In theory, all education is directed toward some "cognitive development," and Hallinan reports no unique benefit from the multicultural content of a course. Further, the observation that students exposed to multiculturalism through course work increase their racial understanding is virtually redundant, akin to observing that students who take mathematics classes learn mathematics. Indeed, the most resilient conclusion from Hallinan's meta-analysis of college studies is that there is no evidence suggesting that diversity *impairs* the achievement of either White or minority students.[48] Hardly compelling.

The Shape of the River,[49] the famous work by William Bowen and Derek Bok—who served as expert witnesses in *Gratz* and *Grutter*, respectively[50]—used the Andrew W. Mellon Foundation's College and Beyond database to track students who had enrolled in 28 elite colleges and universities, 70% at one of 24 private universities, and the other 30% at one of four large public schools. The study focused on long-term outcomes—students were surveyed first in 1976, and again in 1989—including advanced degree attainment, employment, earnings, job satisfaction, civic participation, and views on race relations.

I rely on the excellent summary by William Kidder[51] to report Bowen and Bok's findings. Several findings relate to interracial sociability. 57% of Blacks and 46% of Whites in the 1976 sample perceived the "ability to work effectively and get along well with people from different races/cultures" as important, with percentages that had improved to 70%

---

46. *Id.* at 748 (discussing Maurianne Adams & Yu-Hui Zhou-McGovern, *The Sociomoral Development of Undergraduates in a "Social Diversity" Course: Developmental Theory, Research, and Instructional Applications* (paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA (1994)).

47. *Id.* at 749 (discussing Thomas R. Bidell et al., *Developing Conceptions of Racism Among Young White Adults in the Context of Cultural Diversity Coursework* (paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA (1994)).

48. *Id.* at 747.

49. WILLIAM G. BOWEN & DEREK BOK, THE SHAPE OF THE RIVER: LONG-TERM CONSEQUENCES OF CONSIDERING RACE IN COLLEGE AND UNIVERSITY ADMISSIONS (1998).

50. Expert report of William Bowen, Gratz v. Bollinger, 122 F. Supp. 2d 811, No. 97-75321 (E.D. Mich. 2000), *available at* http://www.umich.edu/~urel/admissions/legal/expert/bowen.html; expert report of Derek Bok, Grutter v. Bollinger, 16 F. Supp. 2d 797, No. 97-75928 (E.D. Mich. 1998), *available at* http://www.umich.edu/~urel/admissions/legal/expert/bok.html .

51. Kidder, *supra* note 34, at 222–23 & nn.309–22.

of Blacks and 63% of Whites by 1989.[52] The 1989 cohort reported a moderate level of interracial interaction, with 88% of Blacks and 56% of Whites saying they knew well at least two students from the other race.[53] Academically, the study found that 75% of Blacks and 81% of Latinos in the 1989 cohort had graduated within six years from the same college they had entered, and an additional 4% of Blacks and 9% of Latinos had transferred and graduated elsewhere. These rates contrasted with 40% Black and 59% White national graduation rates at large universities.[54] 40% of Blacks in the 1976 cohort had obtained doctorates or professional degrees, compared to only 8% nationally.[55]

The Shape of the River suffers from many methodological defects, defects repeated by the *amici curiae* researchers upon whom the *Grutter* Court principally relied. As such, I defer discussion of the flaws until after I have presented the *amici curiae* research.

C. *Contrary Research*

Many studies have found that increased student body diversity is unrelated to educational outcomes, even associated with negative outcomes. Justice Thomas relied on such research in his *Grutter* dissent.[56] One reviewer[57] lamented the inconsistent findings regarding the effect of desegregation on the academic achievement of primary and secondary students; another study[58] found that diversity neither harmed nor benefited college students' academic performance. Louise Bohr and colleagues[59] found no significant reading, mathematics, or critical thinking test score differences between Black students educated in two historically Black colleges, versus those educated in sixteen predominantly White col-

---

52. *Id.* at 222–23 (quoting BOWEN & BOK, *supra* note 49, at 220–21). Perhaps it is troubling that these rates are not substantially higher than 50%.

53. *Id.* at 223 (citing BOWEN & BOK, *supra* note 49, at 231–34). The likelihood of Whites knowing well at least two Blacks grew with the percentage of Blacks in a student body, and those Whites who knew well at least two Blacks in college were more likely to know at least two Blacks after college. Of course, this first finding sounds like an accident of availability, and I wonder whether the two Blacks that Whites knew well after college were different than the two they knew during college.

54. *Id.* at 222 & nn.309–10 (citing BOWEN & BOK, *supra* note 49, at 56–57 & fig. 3.1).

55. Id. at 222 (citing BOWEN & BOK, *supra* note 49, at 98 fig. 4.2).

56. *Grutter*, 539 U.S. at 364–65 (Thomas, J., dissenting).

57. DAVID ARMOR, FORCED JUSTICE: SCHOOL DESEGREGATION AND THE LAW 59–116 (1995).

58. Harry Holzer & David Neumark, *Assessing Affirmative Action*, 38 J. OF ECON. LITERATURE 483 (2000).

59. Louise Bohr et al., *Do Black Students Learn More at Historically Black or Predominantly White Colleges?* 36 J. OF C. STUDENT DEV. 75, 77–79 (1995).

leges; and another study[60] using the same data set found that the Black students at the historically Black colleges actually scored higher on the writing skills test. A final study[61] found that college retention for both Blacks and Whites *diminished* with increasing campus diversity.

Switching to social outcomes, several studies found that increased student body diversity was correlated with greater conflict between different-raced students. Laboratory studies have shown that contact between different-raced students does not necessarily foster intergroup relationships, and instead often spurs interracial antagonism.[62] Similarly, one study[63] found that increased diversity was associated with increased racial insularity in housing, social activities, customs and beliefs.

Several studies have exposed a distaste for affirmative action and diversity among both college students and faculty. A sizeable survey of 1,600 college students and 2,400 faculty[64] found that higher proportions of Blacks or Latinos in a student body were associated with less satisfaction with one's education and the work ethic of one's peers, as well as more frequent claims of discrimination. 85% of students surveyed, including 71% of minority students, rejected racial preferences in admissions. Another survey[65] found that a majority of 800 university faculty opposed affirmative action. Finally, Black law students appear to value diversity more than do White students.[66]

Of course, much like a positive correlation between diversity and educational achievement does not imply that diversity *caused* such achievement, a negative correlation between diversity and educational outcomes does not imply that diversity is to blame. Nonetheless, contrary research illuminates a diversity *counter*-rationale, or at least creates suspicion of the diversity rationale. If racial diversity breeds interracial understanding and cooperation at one school, is it so hard to believe that it might breed antagonism at another? If you doubt that diversity actually

---

60. Ernest T. Pascarella et al., *Influences on Students' Openness to Diversity and Challenge in the First Year of College*, 67 J. OF HIGHER EDUC. 174 (1996).

61. Mitchell J. Chang, Racial Diversity in Higher Education: Does a Racially Mixed Student Population Affect Educational Outcomes? (1996) (unpublished Ph.D. dissertation, UCLA) (on file with UMI Dissertation Services).

62. LEIGH THOMPSON, THE MIND AND HEART OF THE NEGOTIATOR 188–219 (2000). *See also* Nancy E. Dowd et al., *Diversity Matters: Race, Gender, and Ethnicity in Legal Education*, 15 U. FLA. J.L. & PUB. POL'Y 11, 27 (2003) (survey of 300 University of Florida Law School students).

63. Chang, *supra* note 61.

64. Stanley Rothman et al., *Does Enrollment Diversity Improve University Education?*, 15 INT'L J. OF PUB. OPINION RES. 8 (2003).

65. NATIONAL ASSOCIATION OF SCHOLARS, NATIONAL FACULTY SURVEY REGARDING THE USE OF SEXUAL AND RACIAL PREFERENCES IN HIGHER EDUCATION (1996), *available at* http://www.nas.org/reports/roper/exsum.htm.

66. Dowd et al., *supra* note 62, at 27.

depresses Black students' writing skill, should you not also be skeptical that diversity improves their critical thinking skill?

III. *AMICI CURIAE* BRIEFS INSTRUMENTAL TO THE *GRUTTER* DECISION

The briefs by *amici curiae* in support of the University of Michigan were enormously influential to the *Grutter* decision.[67] In this section, I review the evidence from the two briefs that most swayed the *Grutter* Court in articulating the compelling interest in the educational benefits of diversity. These briefs were submitted by 1) the American Educational Research Association, the Association of American Colleges and Universities, and the American Association for Higher Education[68] [hereinafter, "Education Brief"]; and 2) high-ranking former generals, admirals, and civilian leaders of the United States Military[69] [hereinafter, "Military Brief"]. I analyze the Education Brief in particular detail, preparing for my discussion in the final section of this Note of the theoretical and methodological weaknesses of the research the brief advances. I also outline the arguments advanced by the Military Brief and explain how these arguments are theoretical only, offering no testable hypothesis to demonstrate diversity's educational benefit. Where the diversity interest cannot be verified, it should not be compelling.

A. *Education Brief*

Grutter and *Gratz* relied principally on the Education Brief to conclude that a diverse student body has educational benefits. The brief, in turn, relies on an expert report by Patricia Gurin,[70] a Professor at the

---

67. Neal Devins, *Explaining* Grutter v. Bollinger, 152 U. PA. L. REV. 347, 366–68 (2003) (noting how, across *Grutter* and *Gratz*, 83 *amicus* briefs supported the University, but only 19 supported petitioners; how 124 members of the House of Representatives, 13 Senators, 23 states, and 91 colleges and universities joined briefs supporting the University, but no Congressman, only Florida State, and no college or university supported petitioners); George, *supra* note 27, at 1635 (the diversity rationale "was urged on the Court in the most impassioned terms by people and institutions of enormous prestige and influence.").

68. Brief of American Educational Research Association et al. as Amici Curiae in Support of Respondents, Grutter v. Bollinger, 539 U.S. 306 (2003) (No. 02-241).

69. Consolidated Amicus Brief of Lt. Gen. Julius W. Becton, Jr. et al., Grutter v. Bollinger, 539 U.S. 306 (2003) (No. 02-241).

70. Expert Report of Patricia Gurin, Gratz v. Bollinger, 122 F. Supp. 2d 811 (E.D. Mich. 2000) (No. 97-75321) and Grutter v. Bollinger, 16 F. Supp. 2d 797 (E.D. Mich. 1998) (No. 97-75928), *reprinted in* 5 MICH. J. RACE & L. 363 (1999), and *available at* http://www.umich.edu/~urel/admissions/legal/expert/gurintoc.html (last visited Mar. 10, 2006). For clarity of pinpoint citations, I refer hereinafter to the pagination from the reprint in the *Michigan Journal of Race and the Law*; but said journal did not reprint the five appendices, so I identify such references only by appendix letter ("A" through "E"), and

University of Michigan, submitted into evidence by the Law School in *Grutter* and by the undergraduate College in *Gratz*. The report was prepared specifically in anticipation of the *Grutter* and *Gratz* litigations.

The brief argues that student body diversity promotes learning, democratic values, and civic engagement, and prepares students for a diverse society, workforce, and clientele. The brief also indicates mechanisms whereby diversity supposedly improves the educational experience. Student body diversity is thought to create diverse classrooms that "challenge students to consider alternative viewpoints and to develop tolerance for differences," promoting the development of "critical thinking skills."[71]

Gurin analyzed three sources of data. The largest data set incorporated surveys from 9,300 students at nearly 200 universities, collected by the Cooperative Institutional Research Program conducted at UCLA [hereinafter, "CIRP study"]. Students completed one survey when they entered college in 1985 and follow-up surveys in 1989 and 1994.[72] Second, Gurin analyzed data from the Michigan Student Study [hereinafter, "MSS study"] collected from 1,134 White students and 187 Black students who entered the University of Michigan in 1990. Surveys were administered at the beginning and end of freshmen year, and at the end of sophomore and senior years. Latinos and other racial minorities were omitted from the data collection, because of an insufficient sample size.[73] Finally, to study mechanisms whereby diversity might produce educational benefits, Gurin analyzed data from undergraduates who entered the University of Michigan in 1990 and enrolled during freshman year in a class in the Intergroup Relations, Community, and Conflict Program [hereinafter, "IGRCC study"]. These students participated in a 10-week dialogue group, designed to help students discern differences and similarities between White and Black students' viewpoints on contested issues; examine differences in viewpoint *within* each race; negotiate conflicts arising in the dialogues; and challenge the groups to find bases for coalition and joint action.[74] Students were surveyed first as participants in the MSS study, again after the IGRCC program was completed, and three years

---

pinpoint citations to these appendices can be verified by reference to the report as available on the University of Michigan website.

71. Education Brief, *supra* note 68, at 3.

72. Gurin, *supra* note 70, app. C. The National Association of Scholars, who submitted *amicus* briefs in favor of both Barbara Grutter and Jennifer Gratz, criticized CIRP for denying access to its data set to persons and organizations against affirmative action programs. *See* Peter Schmidt, *Report Questions Michigan's Defense in Affirmative-Action Lawsuit*, CHRON. OF HIGHER EDUC., Apr. 13, 2001, at A37.

73. Gurin, *id.* app. C.

74. *Id.* Course content included contemporary analysis of group inequalities in economic, educational, and political arenas, and associated policies regarding immigration, bilingual education, affirmative action, sexual harassment, and Middle East peace initiatives.

later at graduation. The study employed a matched-sample control group of students not in the IGRCC program.[75]

Gurin conceptualized three forms of diversity as independent variables. First, she used "structural diversity," the mathematical racial composition of a school's student body. Structural diversity theoretically produces institutional transformations that create the possibility for "classroom diversity"—the incorporation of knowledge about diverse groups into the curriculum—and "informal interactional diversity," the opportunity to interact on campus with students from diverse backgrounds. Although structural diversity makes the latter two varieties possible, Gurin credits only classroom and interactional diversity with producing educational benefits.[76]

Synthesizing the results from the three studies, Gurin concludes that diversity produces both "learning" and "democracy" outcomes.[77] Students exposed to the most diversity in classroom settings had the "greatest engagement in active thinking processes, growth in intellectual engagement and motivation, and growth in intellectual and academic skills."[78] They employed "conscious, effortful, deep thinking," in contrast to thinking that is preconditioned or stereotyped, enabling them to better understand others' perspectives and manage conflicts.[79] During college, students exposed to greater diversity were more engaged as citizens and interacted more with people from different races and cultures. This engagement and interaction extended to the five years immediately following college, when students exposed to greater campus diversity were more likely to participate in activities serving community and promoting racial understanding, and were more likely to have cross-racial friendships.[80]

Some students in the IGRCC study claimed that their undergraduate education affected the way they thought about diversity: 40% of Blacks, 25% of Asians, and 28% of White students indicated that some course at Michigan had significantly impacted their views on diversity, and for 95% of these students, the impact was positive.[81] Unfortunately, Gurin omits discussion of the troubling fact that, in the IGRCC study, large majorities of all students apparently indicated that their views had *not* been affected by classroom diversity. Further, at least one study employing the same CIRP database as Gurin found that student racial diversity did *not* affect academic outcomes including grades, dropout rates,

---

75. *Id.* Gurin does not identify how many students participated in the IGRCC study.

76. Gurin, 5 MICH. J. RACE & L. at 376–77.

77. *Id.* at 365–66.

78. *Id.* at 365.

79. *Id.* at 372.

80. *Id.* at 366.

81. Gurin, *supra* note 70, app. E.

and performance on seven standardized tests. Any weak effects were mediated and indirect.[82]

Gurin's report focuses exclusively on undergraduate education. The Education Brief also discusses, with less detail, three studies on the educational benefits of diversity in law school. The first consisted of survey data from 1,820 students at Harvard and Michigan Law Schools, collected with an 81% response rate using the Gallup Poll, as well as e-mail surveys from several other elite law schools (with lower response rates).[83] Two-thirds of the students were White. Majorities of students of all races at each school said diversity enhanced their thinking about problems and solutions, the way topics were discussed in classes and outside the classroom, and their ability to work effectively and get along with members of other races. Conflicts because of racial differences challenged them to rethink their values, including views regarding the equity of the criminal justice system, property rights, and contractual rights.[84] Overall, 90% of students from each school reported an overall positive impact of diversity on their educational experience; and fewer than 1% said that having diverse peers had a negative impact on their education.[85]

The second law school-specific study was a survey of more than 2,000 Michigan Law School alumni (half of whom were racial minorities) graduating between 1970 and 1996.[86] The Education Brief highlights three findings from the study. First, large proportions of alumni placed considerable value on the contribution diversity made to their classroom experiences in law school. Second, twice as many (50%) White male alumni who had graduated in the 1990s responded positively to diversity than did White males graduating in each of the previous two decades (25%), when there were fewer non-White students enrolled. Despite the growth, such statistics remain discouragingly low. Third, the survey found that minority alumni were more likely than other alumni to engage in government and public interest work and to serve individuals of their own race or ethnicity.[87] Of course, this last finding suggests that minorities educated alongside diverse peers return to a racially insular environment upon graduation, hardly laudatory of diversity's benefit.

---

82. ALEXANDER W. ASTIN, WHAT MATTERS IN COLLEGE?: FOUR CRITICAL YEARS REVISITED 186–244 (1993).

83. Gary Orfield & Dean Whitla, *Diversity and Legal Education: Student Experiences in Leading Law Schools, in* DIVERSITY CHALLENGED: EVIDENCE OF THE IMPACT OF AFFIRMATIVE ACTION 143 (Gary Orfield & Michal Kurlaender eds., 2001). The study is discussed in the Education Brief, *supra* note 68, at 16–18, 21. The survey is summarized by Kidder, *supra* note 34, at 226–27 & nn.352–60.

84. Orfield & Whitla, *id.*

85. *Id.* at 160–61 (*discussed in* the Education Brief, *supra* note 68, at 21).

86. Richard O. Lempert et al., *Michigan's Minority Graduates in Practice: The River Runs Through Law Schools*, 25 L. & SOC. INQUIRY 395 (2000). The study is discussed in the Education Brief, *supra* note 68, at 18–19, 21.

87. *Id.*

The final law school-specific study reported in the Education Brief, a survey of 558 law school faculty, was conducted by the Association of American Law Schools in 1999.[88] Faculty supported student body diversity, believing that diversity helps students confront racial stereotypes by broadening the variety of experiences shared in the classroom. 75% of faculty felt strongly that having a diverse student body is important to their law school's mission.

B. *Military Brief*

The Military Brief caught much of the Justices' attention at oral argument for *Grutter*.[89] The brief portrays student body diversity as a compelling government interest necessary (and therefore subsidiary) to its paramount interest in national security: "*Amici* submit that the government's compelling interest in promoting racial diversity in higher education is buttressed by its compelling national security interest in a cohesive military[,]" for, " '[i]t is obvious and unarguable that no governmental interest is more compelling than the security of the Nation.' "[90] Of course, national security was never an interest advanced by the defendant Law School.

Nonetheless, the Military Brief describes how racial tensions in the armed services during the Vietnam War led to violence in the ranks, impairing the ability to fight.[91] The tensions created racial polarization and disciplinary problems, and impeded the flow of information through the chain of command. According to the brief, the major impediment to order was the lack of minority officers: the armed forces had become a mix of racially diverse enlisted ranks, commanded by an overwhelmingly White officer corps. For example, in 1962, only 1.6% of all commissioned officers were Black. The brief argues that better integration of the ranks requires better integration of the officer corps.

In pursuit of numerical integration, the military uses affirmative action in admissions to its own academies and to the Reserve Officers' Training Corps (ROTC) scholarship program. Because racial minorities from highly selective colleges achieve significantly better officer reviews

---

88. RICHARD A. WHITE, LAW SCHOOL FACULTY VIEWS ON DIVERSITY IN THE CLASSROOM AND THE LAW SCHOOL COMMUNITY (May 2000), *available at* http://www.aals.org/statistics/diverse3.pdf. The study is discussed in the Education Brief, *supra* note 68, at 19–20.

89. *See* Transcript of Oral Argument, Grutter v. Bollinger, 539 U.S. 306 (2003) (No. 02-241). Several Justices call on counsel for the petitioner to explain their position that diversity is not a compelling state interest given the claims put forth in the Military Brief. *Id.* at 7–17, 19–22.

90. Military Brief, *supra* note 69, at 7–8 (quoting Haig v. Agee, 453 U.S. 280, 307 (1981)).

91. *Id.* at 6–7, 14–17.

than minorities from less selective colleges, the brief argues that increasing the number of minority students at selective colleges is essential to increasing the supply of minority officer candidates.[92] In other words, selective colleges feed the military with intelligent, diverse officers. The brief argues, "At present, the military cannot achieve an officer corps that is *both* highly qualified *and* racially diverse unless the service academies and the ROTC use limited race-conscious recruiting and admissions policies."[93]

The Army, Navy, and Air Force all use race as a factor in recruiting and admissions.[94] The service academies and ROTC set goals for minority officer candidates, which they pursue by financial and tutorial assistance and tailored recruitment programs. West Point sets yearly targets for minority admissions, aiming to "'represent the society we come from.'"[95] The Naval Academy targets 7% Blacks and 4% Latinos in its student body, aiming to achieve the Naval directive of 6% Blacks and 3% Latinos in its officer inventory. The Air Force Academy does not employ specific targets, but does give racial preferences: from 1991–1995, 76% of qualified minority applicants received offers, compared to 51% of White applicants. Finally, the ROTC, which produces 48% of active duty officers, offers a substantial number of scholarships at historically Black colleges and at institutions with high Latino enrollment; and each service's ROTC program works to meet its service's goals for commissioning minority officers. Today, the ranks are better integrated, with almost 40% minority servicemen and women, including 21.7% Black, 9.6% Latino, 4% Asian, and 1.2% Native American. 19% of officers are minorities (including 8.8% Black), an improvement on the past but still proportionately low relative to minority representation in the ranks.[96]

At baseline, the military's articulated interest in campus diversity is motivated only by its interest in having a diverse, qualified applicant pool available for military service. Elite schools provide the military with a recruiting pool pre-selected for intelligence. If the feeder schools are also racially diverse, then other, more difficult and costly recruiting efforts need not be undertaken. But other methods *are* available: the military could expand the scope of its recruitment effort outside of universities, or increase the frequency of promoting qualified minority candidates from within; or it could beseech Congress to enact a compulsory service

---

92. *Id.* at 27.

93. *Id.* at 5 (*quoted by Grutter*, 539 U.S. at 334) (emphasis in original).

94. *Id.* at 7, 19–21, 25, 35–36.

95. *Id.* at 19 (quoting Adam Clymer, *Service Academies Defend Use of Race in their Admissions Policies*, N.Y. TIMES, Jan. 28, 2003, at A17).

96. *Id.* at 12, 17 (citing U.S. DEP'T OF DEF., STATISTICAL SERIES PAMPHLET NO. 02-5, SEMIANNUAL RACE/ETHNIC/GENDER PROFILE BY SERVICE/RANK OF THE DEPARTMENT OF DEFENSE & COAST GUARD 4 (2002)).

program. Relying on intelligent, diverse student bodies at selective colleges is convenient, but by no means necessary.

Further, the military is *un*interested in student body diversity in and of itself, or in the educational benefits diversity may yield. Indeed, the military offers no proof that the diversity at elite schools *produces* qualified officers.[97] Diversity at the college level is useful only as a means to a recruiting end. Diverse, intelligent student bodies are constructed by admitting diverse, intelligent students. Diversity does not create the students' intelligence. The military's argument concerns access, not pedagogy.

The racial preferences that the service academies grant to minorities may themselves be legally justified as a particular remedy for specific, prior discrimination by the military and its academies.[98] But the *Grutter* Court held that student body diversity confers educational benefits.[99] Minority representation in the military is not an educational benefit. Furthermore, diversification of the ranks and officer corps may be a compelling interest when such diversity enhances the functioning of the military and, therefore, national security.[100] However, much like the narrow remedial interest, the interest in national security would justify only affirmative action practiced directly by the military through recruitment, enlistment and promotion; or affirmative action practiced directly by the service academies. This interest is also narrow. Where the brief errs is in extrapolating the interest to all of higher education. The goal of diversifying the military cannot justify racial preferences granted by all public universities—granted on the military's behalf—when the fraction of university students entering an ROTC program, and the fraction of graduates ultimately entering the military, are necessarily so small.

---

97. The Military Brief is void of any attempt to empirically measure the relationship between diversity and educational benefit, relying only on assertions. For instance, while the brief claims that diverse campuses are helpful toward educating future officers to work with and lead diverse ranks, its authors cite no evidence that diverse campuses actually have this effect or provide this training. Military Brief, *supra* note 69, at 28.

98. *See supra* notes 11–12 and accompanying text.

99. *See supra* notes 15 and 18.

100. Even this interest is doubtful. The argument relies on an assumption from 1960s and 1970s data that, absent a diverse corps, racial tensions would resurface, destroying unity and impeding the ability to fight. *See supra* text accompanying note 91. However, the military offers no evidence that this assumption is valid today, and it is plausible that the tension, if any, would not be so intense, or that it might be better managed. Today, the percentage of officers who are minorities (19%) remains less than half the percentage of servicemen and women who are minorities (40%), *supra* note 96, yet racial tensions do not seem to plague a United States military heavily engaged across the world. Is it possible that only 15% or 10% minority officers would do the trick today? Findings from social science, after all, are context-dependent and, often, transient.

IV. DIVERSITY RATIONALE: PROBLEMS OF METHOD AND THEORY

Paul Brest, former Dean of Stanford Law School, advances a theory of why a diverse classroom is important to legal education:

> American legal education is highly interactive, aimed at getting students to examine the law from all possible perspectives.... Students are pressed to examine even well-settled cases with a skeptical eye. This process is designed to teach problem-solving skills and to impart the critical stance characteristic of all good lawyers. This method of education depends on students bringing diverse perspectives to the classroom.[101]

In this final section, I discuss the many flaws in the theory and research linking a facially diverse student body to educational benefits. Many of the flaws are methodological; by examining closely the research advanced by the Education Brief, I expose the dubious assumptions made and errors committed. Other difficulties exist in how the diversity rationale is explained by the Education Brief and by the Court. These difficulties are broad: can "diversity" be measured? Can its effect on educational outcomes be modeled and proven? Is that effect reliable and universal? Can universities meaningfully harness diversity's educational power?

My answers to these questions are, "No." The diversity rationale, as applied by the University of Michigan Law School [hereinafter, "Law School"] and upheld in *Grutter*, suffers from its simultaneous wish to be broad and its required narrowness under the microscope of strict scrutiny. This dichotomy injures the logic. Still, even assuming a lucid theory open to empirical testing, the research presented in the Education Brief provided a woefully inadequate test. The Court was wrong to accept a theory that was not, and cannot be, validated.

For clarity, I categorize the weaknesses below, but one should not underestimate how much the categories intersect and the defects interact. The first four criticisms address the difficulty in measuring abstract variables like "diversity" and "educational benefit," and are followed by an explanation why the relationships between these variables cannot be validly determined or generalized. I then discuss several prevalent biases in the research, and I conclude with two demonstrations of how the cheerleaders of diversity are not fully committed to the rationale they advance.

---

101. Paul Brest, *Some Comments on* Grutter v. Bollinger, 51 DRAKE L. REV. 683, 683–84 (2003).

A. *Mismeasurement of the Independent Variable, "Diversity"*

To test how diversity affects educational outcomes, it is critical first to define and measure diversity appropriately and precisely. The Law School's admissions policy professes to seek multiple varieties of diversity, not just racial diversity.[102] It is this broad sense of diversity that the *Grutter* Court requires; but it is this same breadth that renders the interest unworkable. In the words of one federal court, "[H]ow can a State dole out race-based preferences when it cannot adequately define the compelling interest it seeks to serve?"[103] The Court pronounces that diversity is not an end in itself, but a means to educational benefit. Too many admissions committees, including the Law School's, interchange these two conceptions.

Even assuming there is a compelling interest in strictly racial diversity (both *Bakke* and *Grutter* insist there is not), the question arises: diverse by what standard? Mapping racial composition of a student body onto composition of the entire country would produce a very different result than a system reflecting regional variation.[104] The Law School's admissions documents refer to 10–17% of total enrollees as a desired goal for African American, Latino, and Native American students;[105] in the past, classes having these percentages supposedly helped the Law School achieve "the kinds of benefits that we associate with racial and ethnic diversity."[106] These numbers are arbitrary and lead to further conundrums: should a biracial student qualify as one-half of a diversity "unit," or a category unto himself? Do you force identification with the race of one parent and not

---

102. The policy explains, "[T]here are many possible bases for diversity admissions," and special consideration might be given for "an Olympic gold medal, a Ph.D. in physics, the attainment of age 50 in a class that otherwise lacked anyone over 30, or the experience of having been a Vietnamese boat person." Grutter v. Bollinger, 288 F.3d 732, 736 (6th Cir. 2002). But the admissions committee made no effort to monitor these types of diversity, operating only on an *ad hoc* basis, if at all. In contrast, every day the Law School admissions director monitored admissions reports classified by race. Grutter v. Bollinger, 137 F. Supp. 2d 821, 842 (E.D. Mich. 2001).

103. Tracy v. Bd. of Regents of the Univ. of Ga., 59 F. Supp. 2d 1314, 1322 (S.D. Ga. 1999).

104. For example, year 2000 Census data show that Michigan's racial composition differs from that of the nation, consisting of greater proportions of Whites (80.2% in Michigan, 75.1% across the United States) and Blacks (14.2% versus 12.3%); but a much smaller proportion of Latinos (3.3% versus 12.5%). http://quickfacts.census.gov/qfd/states/26000.html (last visited Jan. 25, 2006).

105. *Grutter*, 137 F. Supp. 2d at 840–41. It is unclear why the Law School limits its preferences to these three groups. Martin Carcieri argues, "Since diversity/critical mass is so crucial, substantial preferences for three races are pronounced narrowly tailored to advance these ends." Martin D. Carcieri, *The Sixth Circuit and* Grutter v. Bollinger*: Diversity and Distortion*, 7 TEX. REV. L. & POL., Fall 2002, at 127, 142 (citation omitted).

106. *Grutter*, 137 F. Supp. 2d at 835 (quoting a draft of the Law School's admissions policy).

the other, or one grandparent and not the other three?[107] When we once again introduce the infinite types of non-racial diversity, the complexity increases exponentially, and "diversity" is reduced to a "buzzword." John Bunzel writes, "Diversity has become a universal good presumed to be so self-evident that it need never be defined or can conveniently be redefined according to the occasion."[108]

Empirically, diversity has been inconsistently measured. One logical measure is the racial composition of a school. A researcher could, for instance, measure the correlation between any outcome variable and the percentage of Black students at a school.[109] The difficulty with this approach is that racial diversity (let alone diversity in a more general sense) includes much more than "Blackness."[110] If a diversity of viewpoints is important to a well-rounded education, using Blackness alone as a proxy for diversity is grossly inadequate.

A broader way to measure diversity, like that employed by the CIRP study and by most of the studies reviewed by Hallinan,[111] is the percentage of students of color at a school, measured collectively. However, such a measure is insensitive to differences in viewpoint between and within racial groups. The imprecision is exacerbated where a study employs a bivariate measure (e.g., Hallinan's majority-White versus minority-White schools) that ignores how increasing diversity incrementally affects the educational experience. A bivariate measure would distinguish meaninglessly between a 1,000-student school with 501 White students and one with 499. To Gurin's credit, she measured campus racial diversity—what she called "structural diversity"—continuously. Even so, any measure of

---

107. If so, you create the potential for gaming the system or lying about race to exploit preferences. *See* Barbara Lauriat, *Trump Card or Trouble? The Diversity Rationale in Law and Education*, 83 B.U. L. REV. 1171, 1202 (2003) (arguing that admissions committees would need to "create specific criteria for each group and require supporting proof of group membership," which would require accessing and verifying family trees).

108. Bunzel, *supra* note 3, at 490; *id.* at 498 ("The term 'diversity' has become a code word that fails to define precisely what it allegedly exalts and what exactly is to be accomplished by those who extol its virtues.") (citation omitted). Another commentator wonders why there is no push to diversify university *faculty* to incorporate various racial, political, and religious viewpoints underrepresented at elite universities. George, *supra* note 27, at 1636.

109. This was the measure employed by Rothman et al., *supra* note 64. As separate independent variables, the authors used the percentage of Black students, of Latino students, and of Asian American students.

110. *See, e.g.*, *Bakke*, 438 U.S. at 315 ("The diversity that furthers a compelling state interest encompasses a far broader array of qualifications and characteristics of which racial or ethnic origin is but a single though important element. Petitioner's special admissions program, focused *solely* on ethnic diversity, would hinder rather than further attainment of genuine diversity.") (emphasis in original).

111. Hallinan, *supra* note 36, at 740–41.

sheerly numerical diversity ignores how a school actually harnesses its diversity to yield educational effects.[112]

Still, any measure of institutional diversity is preferable to the ridiculous classroom diversity measure that Gurin used in the CIRP and MSS studies. In the CIRP analysis, a student's exposure to classroom diversity was a bivariate measure of whether she had ever enrolled in an ethnic studies course;[113] and the thrust of the MSS classroom diversity index was a student's senior year self-report of whether he had ever taken a course that significantly impacted his views on racial diversity and multiculturalism.[114]

These measures are not only misleading, but potentially offensive in their assumptions. The CIRP measure assumes that ethnic studies courses are never passive lectures, but always filled with debate and the sharing of diverse viewpoints. It assumes some degree of the stigma and spokesmanship Justice O'Connor has been otherwise careful to avoid.[115] The measure also assumes that students enrolled in ethnic studies courses have automatically been exposed to "diversity." Thus, by Gurin's analysis, twenty White students sitting alone in a classroom have enjoyed a diverse experience—nay, a diverse *educational career*—if the course is one in ethnic studies. Why, then, were racial admissions preferences ever necessary?

The MSS measure is similarly imprecise, as it includes no quantitative, continuous component. A typical college student will take 30, 40, or more courses across a four-year college career. How is enrollment in *one* ethnic studies course a meaningful metric for a student's exposure to classroom diversity across her entire college career? If God treated Noah to one day of sunshine, it would be absurd to suggest that Noah enjoyed a spell of good weather. At the very least, one should consider how much rainfall Noah endured on the other 39 days and nights. Likewise, if a student takes one course in race relations and 39 courses in Classics, it would be absurd to accept that the student had a diverse education.

In sum, when diversity is measured too generally (e.g., by the proportion of minorities in a student body), no attention is paid to how individual students interact, or how an institution manages the education of its students to realize the purported educational benefits of those students' diversity. Although structural diversity on campus as a whole might set the stage for more particularized classroom and interactive diversity, only the latter two types are theorized to produce educational benefits, and Gurin presents no evidence that structural diversity actually has such a trickle-down effect.[116] In contrast, when diversity is measured too nar-

112. *See infra* Part IV.I.

113. Gurin, *supra* note 70, app. C and 5 MICH. J. RACE & L. at 382.

114. *Id.*

115. *Grutter*, 539 U.S. at 319–20; *Croson*, 488 U.S. at 493; *supra* note 13.

116. The National Association of Scholars has accused Gurin of being unable to show clear educational benefits of racial diversity, causing her to exaggerate the benefits of

rowly (e.g., by enrollment in ethnic studies courses), researchers risk studying the effects of a proxy far removed from the underlying construct it seeks to represent. Absent a showing that enrollment in such courses is meaningfully correlated with any "true" measure of campus diversity, the proxy is of shallow, if any, worth.

B. *Race as an Imperfect Proxy for Viewpoint*

A compelling diversity interest logically applies to all brands of diversity, not just to a limited number of groups;[117] but in practice, the Law School's admissions committee (like most) focuses almost exclusively on race.[118] If the goal is intellectual diversity, there may be a better case for privileging factors like age, work experience, or country of origin over race.[119] Moreover, exposure to different cultures, viewpoints, and experiences may sometimes best be served by privileging a White applicant over a non-White applicant.

There is wide disagreement about how well racial diversity serves as a proxy for viewpoint diversity,[120] but the proxy is inescapably imprecise. There are at least two reasons why the Law School—and the Court—has accepted imprecision over a more direct measure of viewpoints. First, direct measurement would present too great an administrative burden. Applicants would need to share their "viewpoints" by a personal interview or written essay, and admissions officers would need to meaningfully categorize those viewpoints into pools from which an admissions

---

"diversity activities," none of which require a diverse student body to occur. Schmidt, *supra* note 72.

117. *Bakke*, 438 U.S. at 314 (ethnic diversity is "only one element in a range of factors" relevant to attaining a diverse student body). Justice Powell argued that racial diversity alone is not compelling, *per se*, and greater breadth is required. *Id.* at 315, 317–18 (an admissions program must be "flexible enough to consider all pertinent elements of diversity in light of the particular qualifications of each applicant, and to place them on the same footing for consideration . . . .").

118. *See supra* note 102.

119. Bunzel, *supra* note 3, at 500; Lauriat, *supra* note 107, at 1184 & n.99 (arguing that increasing the number of international students rather than domestic minority students would be more effective in preparing students for a global marketplace).

120. *E.g.*, *Grutter*, 288 F.3d 732, 789 (Boggs, J., dissenting) ("[T]he nature and benefits of the *experiential* 'diversity' that the Law School claims ultimately to seek is conceptually disconnected from the *racial and ethnic* diversity that it primarily seeks.") (emphasis in original); Bunzel, *supra* note 3, at 507 ("[T]here is wide disagreement over whether more racial diversity, and therefore race preferences in admissions, is the fairest or most efficient way to achieve increased viewpoint diversity."); Carcieri, *supra* note 105, at 134 (explaining how the Law School's limitation of preferences to African Americans, Latinos, and Native Americans is a "both over- and under-inclusive means of obtaining diversity of experience."); Note, *An Evidentiary Framework for Diversity as a Compelling Interest in Higher Education*, 109 HARV. L. REV. 1357, 1366–69 (1996) [hereinafter, "*Evidentiary Framework*"] (providing a concise summary of how race associates with experience).

committee could draw. The system would require faith that applicants were sharing viewpoints honestly held, and the assumption that students even have clearly developed views in the first place. At an extreme, viewpoints of matriculated students would need to be monitored over time, and students would risk being dismissed if their viewpoints changed.[121]

The second reason why the Law School uses race as a proxy for viewpoint is that, given that all viewpoints are equally valid under the First Amendment,[122] a genuine solicitation of representative viewpoints would require the school to grant a preference even to marginal opinions, even to viewpoints it hates.[123] After all, if the Law School wishes to enhance the education it provides by exposing students to a broad array of viewpoints, excluding any particular viewpoint might be counterproductive. Instead, the Law School chooses to grant preferences only with an eye extremely sensitive to race. The *Grutter* Court, in looking the other way, hastily abandoned its own, broad diversity requirement.

C. *Inconsistent Definitions of "Critical Mass"*

The Law School enrolls a "critical mass" of minority students to avoid subjecting them to stigma, and to diminish stereotypes of a "minority viewpoint."[124] Unfortunately, "critical mass" seems impossible to define concretely without resort to any poisonous "quota." The Court, relying on testimony by Law School Dean Jeffrey Lehman, defined "critical mass" as "numbers such that underrepresented minority students do not feel isolated or like spokespersons for their race."[125] However, making "critical mass" depend on the feelings of underrepresented minorities requires that the definition be variable. A solitary but extraordinarily strong-willed minority person may constitute critical mass at one institution, where critical mass might require 1,000 weaker-willed persons on another campus. The Education Brief also uses the term "critical mass" nondefinitively: "[W]hen an institution such as the Law School has acted to admit a critical mass of minority students, it also strives to admit enough students to represent varied viewpoints and perspective *within* underrepre-

---

121. *See* Carcieri, *supra* note 105, at 136 ("This is the wild goose chase and social engineering nightmare to which [the University of Michigan's] claim to seek viewpoint diversity logically leads.").

122. *E.g.*, R.A.V. v. City of St. Paul, Minn., 505 U.S. 377, 430–31 (1992).

123. Yale College, for instance, recently admitted a former Taliban diplomat to a nondegree special student program, and was considering his application to a regular degree program at the time this Note was published. This decision has been the source of much controversy. *See generally* Chip Brown, *The Freshman*, N.Y. TIMES MAG., Feb. 26, 2006, at 55.

124. *Grutter*, 539 U.S. at 320 (quoting App. to Pet. for Cert., at 215a). Notice the assumption that some non-minority students believe all minorities share identical views. Lauriat, *supra* note 107, at 1198.

125. *Grutter*, 539 U.S. at 318–19.

sented groups."[126] This is an important point, but there is no evidence that the Law School is actually working to ensure that within-group viewpoint diversity is attained.

Exacerbating the impossible ambiguity is the fact that the mass that is critical apparently varies by minority. Chief Justice Rehnquist's *Grutter* dissent points to data from 1995 to 2000 showing tight correspondence between a given race's representation in the Law School's applicant and admittance pools. The percentages of African American applicants and admittances differed on average by only 0.4% each year; percentages of Hispanics differed by 0.5%; and percentages of Native Americans differed by 0.3%. However, the absolute number of these admitted minorities differed greatly across the groups: an average of 100 African Americans, 52 Hispanics, and only 15 Native Americans were admitted per year.[127] If a "critical mass" of a particular minority is required for students to appreciate that minority's viewpoint(s) without stigma, one would expect equal-sized "critical masses" for each race, rather than mapping onto proportional representation in the broader culture.[128] It makes no sense that only 15 Native Americans are enough to defeat stigma, while 52 Hispanics and 100 African Americans are needed for the same effect.[129]

D. *Mismeasurement of the Dependent Variable, "Educational Benefit"*

1. Questionable "Benefits"

Proponents of the diversity rationale claim that diversity improves such intangible markers of achievement as cognitive development, creativity, study habits, attitudes and beliefs, and the capacities to understand and appreciate others' perspectives.[130] Unfortunately, these outcomes are not objectively measurable like standardized test scores or grades, so they inevitably require assumptions by researchers as to what constitutes a valuable educational benefit. Indeed, by focusing on intangible

126. Education Brief, *supra* note 68, at 25.

127. *Grutter*, 539 U.S. at 383–84 (Rehnquist, C.J., dissenting) (Tables 1, 2, and 3).

128. George, *supra* note 27, at 1638–39.

129. *See* Peter Kirsanow, *Michigan Impossible*, NAT'L REV. ONLINE (July 1, 2003), *available at* http://www.nationalreview.com/comment/comment-Kirsanow070103.asp (on file with the Columbia Law Review).

130. Education Brief, *supra* note 68, at 11 (citing Supplemental Expert Report of Patricia Y. Gurin, Grutter v. Bollinger, No. 97-75928 (E.D. Mich. 2001)) (arguing that objective criteria such as grades, GRE scores, and graduate school admissions fail to explore "active, complex thinking or intellectual engagement"; that few meaningful comparisons can be made between students by using such measures; and that such measures fail to demonstrate improvement over baseline knowledge); Hallinan, *supra* note 36, at 740–41.

educational outcomes, the diversity theory assures its own incompatibility with empiricism, and asks the Court for its faith.

Subjective measures of achievement pervade Gurin's research. For example, students self-rated their abilities compared to the average person their age; their aspirations in the areas of motivation, citizenship, and engagement; and their growth since college.[131] Such measures are entirely relative and suffer from the self-report bias described below.[132] Other measures seem impossibly vague, included primarily as buzzwords. For instance, the Education Brief identifies "civic preparedness" as the central mission of higher education;[133] but it would be difficult to find even two educators or politicians who would agree as to what such preparedness would entail.

Finally, it is questionable whether some outcomes, if genuinely produced by diversity, are actually beneficial. Too often, the research in the Education Brief focuses on *impact*, not benefit. For instance, the Michigan and Harvard Law Schools study asked students whether they thought their discussions with diverse peers would "impact" the legal and community issues they would encounter as professionals,[134] with no identification of whether such impact would steer students toward more, or less, multicultural involvement. Impact is not synonymous with benefit, nor are "benefits" universal commodities to all intended beneficiaries. For instance, "preparing students for a global marketplace" would seem of greater consequence to a would-be businessman or diplomat than a would-be physicist or poet.

Sometimes, what is portrayed as a benefit may be framed otherwise. For example, in the MSS study, the percentage of White students with at least one close non-White friend increased from 32% to 46% during their time at Michigan, and the proportion of Black students with at least one close non-Black friend increased from 47% to 54%. While these increases may properly be deemed benefits, is it not troubling that about half of all students lacked a *single* different-raced friend after a full four years in a diverse student body? Delving further, only 29% of White students and 25% of Black students said they "shared personal feelings and problems" with students of the other race; 23% of Black students said their relationships with White students were "guarded and cautious," and 15% felt they were "tense, somewhat hostile."[135] Suddenly, interracial relationships at Michigan appear much more strained.

---

131. Gurin, *supra* note 70, app. C and 5 MICH. J. RACE & L. at 380–82.

132. *See infra* Part IV.D.2.

133. Education Brief, *supra* note 68, at 12.

134. *Id.* at 20.

135. Gurin, *supra* note 70, app. E.

2. Reliance on Self-Reported Data

Gurin's research, the research in *The Shape of the River*, and most university studies of the diversity rationale, all suffer from self-report bias. Whenever data are collected by survey, there is a chance that respondents' self-reported beliefs may not match what they truly believe, and that subjects may answer in a way they believe the researcher desires. Self-report bias may be conscious or mindless. A student may report greater enthusiasm for diversity than he genuinely feels, deliberately trying to portray himself as enlightened to the researcher (and to himself). Thus, some of the 56% of Whites in Bowen and Bok's 1989 cohort who reported knowing well at least two Black persons[136] may have answered dishonestly, wishing not to appear racist. Or a student at a diverse university may implicitly defend diversity more strongly than a student at a homogenous university, trying to reconcile her "commitment" to diversity with her prior choice to attend a diverse school.[137] Even a meticulously honest participant might unconsciously adjust her interpretation of what it means to know a person "well" if she regards herself as person who *should* be friends with different-raced people.

All of Gurin's research for the Education Brief relied on self-reported data. The Education Brief defends her use of self-assessments as a credible and widely accepted practice,[138] and Gurin herself notes that self-reports of learning outcomes are correlated with traditional measures of achievement, including GRE scores and faculty's reports of student growth.[139] Again, however, such objective measures do not speak to the way diversity is theorized to benefit education, and even a very high correlation—say, 0.7—would still leave unexplained over half the variance in each variable.[140] Are we willing to tolerate such doubt in the articulation of a compelling interest justifying discrimination that would otherwise be unconstitutional?

---

136. *See supra* note 53 and accompanying text.

137. The theory of cognitive dissonance suggests that, contrary to the traditional theory that attitudes dictate behavior, a person may also change her attitudes to correspond to her behavior, akin to a rationalization. *See generally* LEON FESTINGER, A THEORY OF COGNITIVE DISSONANCE (1957).

138. Education Brief, *supra* note 68, at 13.

139. Gurin, *supra* note 70, app. C (citations omitted).

140. The square of a correlation r between two variables indicates the proportion of variation in one variable that can be predicted or "explained" by values of the other variable. Supposing (generously) that GRE scores correlate 0.7 with an abstract variable like self-reported appreciation of diverse viewpoints, the self-report measure could still only predict less than half ($0.7^2 = 0.49 = 49\%$) of the variance in GRE scores.

E. *Invalidity*

1. Internal Invalidity: The Inability To Infer Causation

Research on the educational effects of diversity relies on a technique called linear regression, used to analyze the relationships between independent variable "inputs" and outcome variables, after controlling for other variables known or expected to bias the relationship(s) of interest. Researchers never actually *manipulate* any measure of diversity; to do so might be unethical and, perhaps, impossible. As a result, all of the research is correlational, and correlation never implies causation. Even if research could show a *perfect* correspondence (correlation = 1) between the proportion of minority students in a school and students' academic achievement, it is logically impossible to infer that diversity is *causing* the achievement. A third variable—e.g., monetary resources—may explain both effects, "buying" both diversity (through the ability to offer scholarships and recruit from a larger geographic area) and achievement (through better teachers, textbooks, computers, and enrichment programs).[141]

In her review, Hallinan is not blind to the problem: "[W]hile these survey analyses are consistent in demonstrating a desegregation effect, they do not identify the mechanisms that explain the observed relationship."[142] She tries to fill the deficiency by referencing a handful of case studies that she claims "provide some insights into the social processes involved. These studies show that peer influence, role modeling, instructional quality, and educational expectations are factors that transmit the effects of desegregation to student achievement."[143] Gurin was similarly searching for causal mechanisms by undertaking the IGRCC case study. For whatever advantage case studies yield in suggesting potential causal mechanisms, the great cost is the inability to generalize the findings from case studies beyond their narrow circumstances.

The "gold standard" of social scientific research is the randomized, controlled experiment, which permits a researcher to infer causality. When subjects are randomly assigned to treatment and control conditions, one can assume probabilistic equivalence between the conditions, in terms of the background traits that subjects bring into the experiment; thus, any observed effect can validly be attributed to manipulations of the independent variable. However, when subjects are *not* randomly assigned to conditions, for instance, by choosing their own conditions by choosing to enroll at particular universities, the research is vulnerable to self-selection bias. Thus, we cannot determine whether an observed outcome

---

141. *See supra* note 41 and accompanying text.

142. Hallinan, *supra* note 36, at 744. Of course, her use of the word "effect" itself assumes the direction of relationship.

143. *Id.*

is produced by the independent variable of interest, or by the preexisting traits subjects brought to the study.[144]

When a study is controlled, so only the independent variable is allowed to vary between conditions, one can safely conclude that observed effects were caused by the independent variable, not by a covariate. However, when "life"—the ultimate covariate—is in no way controlled during the study, it is impossible for a researcher to isolate any effect caused by student body diversity. We see this in the IGRCC and in Hallinan's cooperative learning program, where one cannot determine whether students' academic achievements upon exiting the program were due to the diversity of their peers, to the sheer strength of cooperative learning, or simply to students' maturation over the course of the programs.

Cause and effect are often blurred. Consider six variables from the CIRP study: "Discussed racial issues;" "Attend a racial/cultural awareness workshop;" "Socialized with someone from a different racial/ethnic group;" "Number of six closest friends in college who were of respondent's race/ethnicity;" "Extent of involvement at Michigan with groups and activities reflecting other cultural/ethnic backgrounds;" and "Number of five multiethnic campus events attended."[145] These measures might be regarded as "benefits" of diversity. That is, we might hope that a person exposed to more diverse students would be more inclined to socialize with other-race students, to discuss racial issues with them, to consider them close friends, and to develop an interest in multiethnic issues. Yet, these were the six measures of interactional diversity that Gurin employed as *explanatory* ("independent") variables in her models. Where a host of variables are measured simultaneously, it is impossible to gauge which variables are explaining which other variables, or if there is even any explanatory power in the first place. Until some condition is *introduced* to one randomized group, and not to the other, any exposure effects cannot be inferred.

I discovered only one randomized, controlled study of the effects of diversity on education outcomes.[146] The authors randomly assigned 357 White college students to small-group discussions, consisting of three subjects and one research collaborator. 123 subjects were in groups where the collaborator disagreed with everyone, 108 were in groups where the collaborator agreed with everyone, and the remaining 126 were in mixed groups.[147] The collaborator was either Black or White. Students wrote

144. See *infra* Part IV.F.1.

145. Gurin, *supra* note 70, app. C.

146. Anthony L. Antonio et al., *Effects of Racial Diversity on Complex Thinking in College Students*, 15 PSYCHOL. SCI. 507 (2004).

147. *Id.* at 508. Unfortunately, 31 different collaborators were used, meaning that a subject's discussion experience may have been impacted not just by her group's agreement/disagreement and racial composition, but also by the particular characteristics of her

essays on a social issue after participating in the discussion groups, and judges rated those essays for "integrative complexity."[148]

Judges rated students' essays higher when the students had participated in a group where the collaborator held a minority opinion. In other words, students who were exposed to more diverse viewpoints in their discussion groups demonstrated greater assimilation of those views in their post-discussion essays. The finding is dubious, however, because the researchers did not control for *pre*-discussion integrative complexity. That is, they did not examine *growth* in integrative complexity from pre- to post-discussion. Still, even if one generously accepts that exposure to diverse viewpoints did lead to greater integrative complexity, one should still be alarmed by a second finding, that the collaborator's *race* was utterly unrelated to integrative complexity, either directly or by interacting with viewpoint. That is, even if diverse viewpoints produced an educational benefit, race served no proxy function with respect to viewpoint.[149] To the extent that university affirmative action policies grant racial preferences and not viewpoint preferences, the only experimental test in the literature demonstrates that the policies will produce no educational benefit.[150]

Another difficulty plagues the ability to discover genuine educational benefits of diversity. Because exposure to diversity must occur over some reasonably long time period in order to exert its theorized effect, diversity studies are typically conducted over many years. A time effect is always an alternative explanation for a program effect in a longitudinal study. During the time a program is being administered, other social variables change. 1985, when the CIRP study freshmen were first surveyed, was arguably a very different time than 1989, when they were surveyed as seniors, or 1994, when they were surveyed five years out of school.[151] The 1980s and 1990s saw a large increase in diversity on college campuses.[152]

---

collaborator, or the tenor and atmosphere the collaborator created for the group. The existence of this confounding variable makes even this experiment somewhat invalid.

148. "Integrative complexity" referred to cognitive styles that involve "differentiation and integration of multiple perspectives." *Id.* at 508. Judges' ratings of the essays were only moderately reliable, Cronbach's $\alpha = 0.62$.

149. *See supra*, Part IV.B.

150. The Antonio et al. study suffered from three other methodological and explanatory flaws worth footnoting. First, it is difficult to generalize the impact of viewpoint diversity from a single, brief interaction in the laboratory to the many, unquantifiable interactions a student will have across a four-year college career. Second, students apparently were not told to write their essays with "integrative complexity." Thus, some students who "got the point" of the experiment may purposefully have tried to integrate multiple viewpoints into their essays, where students ignorant to the purpose may not have consciously made such effort. Finally, ratings of integrative complexity were very low throughout all conditions, the highest mean rating in a condition being 1.91 on a seven-point scale. Such a restricted range suggests that the ratings were unreliable and invalid.

151. Gurin, *supra* note 70, app. C.

152. Total minority enrollment at the nation's colleges and universities increased by 122% from 1981 to 2001. WILLIAM B. HARVEY, MINORITIES IN HIGHER EDUCATION 2002–

The CIRP students were exposed not to a constant level of diversity, but to increasing heterogeneity over the course of the study. Some participants may, consequently, have mistaken their university's appreciation of diversity partly for their own, or at least had greater opportunity to become sensitive and responsive to diversity. In addition to the admissions trend, in the middle of the CIRP study, there was significant political activity relevant to affirmative action and campus diversity, including enactment of the Civil Rights Act of 1991.[153] The Act, by strengthening existing civil rights laws, may have provided a symbol rallying positive attitudes—or suppressing negative attitudes—toward diversity.

2. External Invalidity: The Inability to Generalize

A goal of most social science—certainly of research intended to bolster a purportedly compelling interest justifying racial discrimination—is to generalize the findings from the participants in any one study to a whole population. For instance, Bowen and Bok found that 70% of Black students in their sample claimed it was important to work effectively with other-race people.[154] To what extent can they conclude that 70% of Blacks at *all* colleges practicing affirmative action will feel similarly? They can only infer this general conclusion from the limited data before them, their precision corresponding to the probability threshold of their inferential statistical test.

One's willingness to forgive potential errors of inference should diminish to the extent that a researcher's sample does not appropriately model the relevant population from which the sample was drawn. Bowen and Bok's student sample provided a poor model of the population of schools practicing affirmative action. They studied 24 private schools but only four public schools, all of them "elite" and selective. But most affirmative action takes place at public institutions,[155] most of which are middle-tier, and some lower-tier, by statistical necessity. Moreover, only affirmative action in public schools (i.e., involving state action) is contested on equal protection grounds, so the focus on private schools is misplaced. *The Shape of the River* is not entirely relevant; that is, valid.

Gurin's research also lacks external validity. The MSS study was conducted only with Black and White students, impairing her ability to generalize the results to groups other than these two races, and ignoring

---

2003: TWENTIETH ANNUAL STATUS REPORT 4–5 (American Council on Education 2003) (analyzing data from the Department of Education's National Center for Education Statistics, the U.S. Census Bureau, and the U.S. Equal Employment Opportunity Commission).

153. Pub. L. No. 102-166, 105 Stat. 1071 (codified in scattered sections of 2, 29, 42 U.S.C.).

154. BOWEN & BOK, *supra* note 49, at 220–21.

155. Michael A. Olivas, *The Shape of the Class*, 24 REV. HIGHER EDUC. 193, 195 (2001).

the infinite other varieties of diversity that the rationale purports to incorporate. Similarly, all three of Gurin's studies—the CIRP, MSS, and IGRCC—used college student samples; yet, in *Grutter*, they were used to defend an affirmative action policy at Michigan *Law School*. There are many reasons why diversity may play a different role in an undergraduate versus law school community. Undergraduate education is typically broad, and exposure to diverse viewpoints may be more critical for students in certain majors (like political science) than to those in other majors (like mathematics). In contrast, a legal education is narrow, intimately tied to policy, and arguably depends more on exposure to diverse viewpoints than undergraduate education. At the same time, a law student is older and presumptively more mature; exposure to diverse viewpoints may be less critical than it is to an undergraduate's comparatively formative years. Finally, law students are probably more likely than undergraduates to have a job or family, or to live and dine off campus, indicating that legal education is less of a life immersion experience than undergraduate education, decreasing the frequency and value of a law student's exposure to diverse peers.

In sum, invalidity is a criticism that can be made of all social science. No study can be completely valid, either internally or externally. This is because, within a single experiment, it is impossible to guard against all potentially confounding variables, and no sample and procedure can perfectly model a population in real life. Insofar as all social science research is at least partially invalid, it is dangerous to mix it with principled constitutional law.

F. *Sampling Bias*

1. Sample Selection

In a number of studies, the researchers' sample selection was biased. In the CIRP study, Gurin excluded from her analyses historically Black colleges, as well as community colleges. At such schools, Gurin "believe[d] that both campus diversity issues and educational processes differ dramatically from those found at predominantly white four-year colleges and universities."[156] She never explains *how* issues and processes might differ; but by excluding historically Black colleges, she dismissed a potentially valuable control group. Her comparisons of racially heterogeneous and homogenously White schools seem to test only the educational effects of diversity on *Whites*—are Whites sincerely the beneficiaries of interest? The exclusion of community colleges from her analyses is less suspect, although the exclusion does limit the generality of her findings.

156. Gurin, *supra* note 70, app. C. Granted, the defendants in *Grutter* and *Gratz* were a predominantly White three-year law school and four-year college.

A more destructive selection bias in diversity research is *self-selection*. Self-selection bias occurs when participants in a study sort themselves into conditions, destroying the randomization that is so valuable to empirical analysis. When participants self-select into conditions, they risk bringing to these conditions systematic differences that remain with them through the period of analysis, ultimately distorting measurement of the outcome variables. The bias manifests itself in a variety of ways.

Unfortunately, *every* survey study comparing students at different schools suffers from self-selection bias, because students choose the schools to which they apply and ultimately attend. For example, if a researcher finds that students at a historically liberal, diverse university choose jobs in urban locations more frequently than do students at a historically conservative, homogenous university, one possible explanation is that students exposed to greater diversity in college seek similar workplace diversity after graduating, a positive effect of campus diversity. However, a more plausible explanation is that students at the liberal university had a pre-existing fondness for diversity before coming to college. The same value they place on diversity prompted them to 1) come to the liberal, diverse university; and 2) choose a diverse urban work force after college. The students who chose the conservative, homogenous university never valued diversity, neither in coming to college nor in pursuing a career. Crediting the universities for creating these differences is too generous.

Conclusions produced by self-selection often sound trivial and self-evident. For instance, Gurin observes, "Students who had taken the most diversity courses and interacted the most with diverse peers during college had the most cross-racial interactions five years after leaving college."[157] Of course they did. They specifically chose to take "diversity" courses and chose diverse friends during college; it would be odd if they suddenly abandoned pursuing diverse relationships after they graduated. Indeed, many of their "cross-racial" friends after college were probably the same "diverse peers" they associated with *in* college. Equally unsurprising is the finding that 95% of students in the IGRCC study who were impacted by diverse viewpoints in their coursework said that impact was positive.[158] The sample consists of students who, at the beginning of their freshmen years, *chose* to enter a course on intergroup relations. They were purely self-selected, and to suggest that the IGRCC Program taught them to value diversity is absurd.

Other examples abound. In the Michigan and Harvard Law Schools study,[159] the students praising diversity for enhancing their educational

157. Gurin, 5 MICH. J. RACE & L. at 409.

158. *See supra* note 81.

159. Orfield & Whitla, *supra* note 83.

experiences were the same students who chose to attend law schools whose admissions policies advertised their commitment to obtaining a diverse student body.[160] Many of these students valued diversity already. Further, the 81% of Michigan and Harvard students who volunteered their time to respond to the survey are unlikely random. Probably, they are a self-selected subgroup having stronger opinions than average about the importance of campus diversity. To the extent that their responses do not reflect those of their entire schools, let alone law students generally, the study is invalid.

By this same reasoning, it is unsurprising that Bowen and Bok, in *The Shape of the River*, found that 70% of Black students and 63% of White students agreed it was important to work effectively with members of the other race, where those students were sampled from 28 "elite" colleges and universities embracive of diversity, as evidenced by their strong affirmative action policies.[161] Again, the findings were survey-generated, meaning that not only was the selection of schools biased, but so too was the sampling of students *within* the schools. Where a survey is voluntary, the students with the strongest opinions on diversity are most likely to participate.

Finally, in one review, Deborah Merritt observes:

> For both minority and white students, attending a more selective college is associated with higher graduation rates and greater earning power. Indeed, African American students who attend selective colleges through affirmative action programs are more likely to graduate from college, obtain advanced degrees, and secure high earnings than are African Americans with similar SAT scores who attend less selective institutions. Thus, minority students capitalize on the opportunities affirmative action programs offer them.[162]

Here, it is difficult to separate an institutional effect from self-selection. The correlations between college selectivity and graduation rate, graduate work, and income potential are unsurprising. Selective colleges would like to explain the correlation as a product of the institutions, but it would be foolish to ignore the fact that what selective colleges select are higher quality, better able and better motivated students.

A half-hearted defense against self-selection bias is statistical control. For instance, when Gurin assessed growth in learning and democracy outcomes related to diversity experiences, she controlled for differences across individuals in their initial positions on learning and democracy

160. After all, one of those schools went so far as to defend its affirmative action policy to the Supreme Court.

161. *See supra* notes 49–52 and accompanying text.

162. Merritt, *supra* note 24, at 1063 (citations omitted).

variables, as well as their relative likelihood to be drawn to intensive diversity experiences. In the IGRCC study, she statistically matched students who did and did not have a "diversity experience."[163] The mechanics of statistical control are beyond the scope of this Note, but the purpose is simply to employ statistical adjustments to "equalize" participants across a host of background variables, attempting to isolate the educational effect directly attributable to diversity; that is, to determine whether classroom and interactional diversity provide any additional explanatory power beyond that explained by background differences. The result, Gurin explains, is a "*less* biased view of the relationship between campus diversity and student outcomes."[164]

Of course, it is impossible to perfectly "match" or "equalize" any two students, to control for all possible background variables. Without random selection of students, combined with random *assignment* of students to research groups—highly unethical in the admissions context—statistical control will always be imperfect, and the best a researcher can hope for is, indeed, "less" bias. Even if students' backgrounds *could* be statistically equalized, growth in one student might mean something different from growth in another. Suppose that in comparing two students' numbers of other-race friends at the end of college, we employ a primitive statistical control by subtracting the number of other-race friends those students had at the beginning of college, thereby controlling for the baseline number of other-race friends. If Student A began college with nine other-race friends and left with ten, and Student B entered college with zero other-race friends but left with one, after applying our statistical control we would discover that the students "grew" equally by one friend. However, it takes an odd set of assumptions to deem this equal numerical growth as subjectively identical. Student A probably entered and left college the same: tolerant and appreciative of diverse friends. One can imagine circumstances whereby Student B's growth was much greater, if, for instance, he were racist or had never met a person of a different race before entering college, but became so open to diversity during college that he now calls a different-race person his *best* friend.

2. Differential Mortality

Another difficulty with longitudinal designs is differential mortality. Whenever subjects are measured at different points in time, there is a risk that some subjects will drop out of the study. In the current context, students completing a first survey may then drop out of school, take a leave of absence, or simply withdraw from participating in the study before a second survey can be administered. Differential mortality is usually not a

163. Gurin, *supra* note 70, app. C.

164. *Id.* (emphasis added).

problem if participants in different conditions drop out at approximately the same rate. However, it becomes a problem when the rates of drop-out are correlated with conditions in the study, or with a third variable of interest.

For example, suppose I track the change in grade point average across the college careers of White students who have taken an ethnic studies course (Group A), versus those who have never taken such a course (Group B). My study finds that students in Group A averaged a 0.5-point GPA increase over the four years, and students in Group B averaged a 0.3-point GPA decrease. One possible explanation is that enrollment in an ethnic studies course somehow enhances academic performance across a student's college career. However, another very different explanation is possible.

Suppose also that 10% of students in Group A dropped out of college during the four years of my study, versus only 2% of students in Group B. Consider what types of students are likely to drop out of college. A variety of causes are plausible, but one obvious candidate is a student performing so poorly as to risk failing out of school. When a student in a longitudinal study drops out of school, he drops out of the study as well. Those participants who remain in the study share something unique: they are able to perform well enough to remain in school without interruption. That is, they have adequate, if not high grade point averages.[165] The fact that 10% of Group A students dropped out of school versus only 2% of Group B students might mean that those students remaining in Group A are more "refined" than those remaining in Group B. While all of A's borderline students may have dropped out, many of B's borderline students may still be "hanging in there," bringing down the whole group's GPA. Then, the condition of enrollment in an ethnic studies course would be unlikely to explain the GPA difference.

Unfortunately, none of the longitudinal studies reported by either Gurin or Hallinan even discuss mortality rates, let alone examine systematic differences between students who completed the study and those who did not.

3. Sample Size and Statistical Power

Statistical power is the odds that a researcher will observe an effect in her particular sample, given that a genuine effect exists in the general population. If there is a real difference between the achievements of students educated in more versus less diverse student bodies, a powerful test using a given sample of students would be likely to detect the difference. Generally, a study is thought to provide a good test if it has power above

165. They share other characteristics as well, not the least of which is the ability to afford tuition.

0.80 (i.e., if there is an 80% chance the test will detect a genuine difference).

Power is related to sample size. The more people a researcher samples, the better the sample approximates the population at large, allowing researchers to make better inferences regarding that population. Suppose I wish to test verbal GRE scores of Black college seniors educated at majority-White versus minority-White universities. Suppose Black students at a given majority-White school genuinely score a mean of 600, and Black students at a given minority-White school genuinely score a mean of 500. In order to have an 80% chance of detecting the difference between Black students' scores at the two schools, I would need a sample of only 17 students from each school.[166] However, if the mean score of Black students at the minority-White school is actually 510 (meaning the true difference between students at the two schools is smaller), I would need to sample 21 students at each school to have the same 80% chance of detecting the genuine difference. If the mean of students at the minority-White school is 570 (meaning the true difference is now down to 30, still quite large), I would need a full 175 students at each school to maintain the same power.

One can see how studies performed with large sample sizes can become quite misleading. In a sense, a study can become *too* powerful, making a big deal out of a small difference. In the above example, sampling 1,570 students from each school would lead to the conclusion that a mean verbal GRE score of 600 is *significantly* different from a mean score of 590—although these scores are only one increment apart on the GRE scale. As another example, with 3,529 students at each school, a difference between a mean IQ of 100 versus a mean IQ of 99 would be judged significantly different.[167]

3,529 students per school (7,058 students total) sounds like a tremendous amount, and it is more students than attend many universities. However, the CIRP study, so central to Gurin's analyses, used over 9,300 students.[168] Dividing this sample in two to make a pairwise comparison between equal samples of 4,650 students would permit the conclusions that a mean IQ difference of 7/8 of one point, or a mean verbal GRE difference of 5.5 points, are statistically significant. Here, statistical

166. This calculation assumes a conventional 95% confidence interval (i.e., a probability < .05 threshold for statistical significance), and a standard deviation of 100 points for each group (the standard for which the GRE aims). Several internet sites permit you to perform power analyses. *See, e.g.*, http://calculators.stat.ucla.edu/powercalc/ (last visited Jan. 25, 2006).

167. This calculation assumes a conventional 95% confidence interval and a standard deviation of 15 points for each group (the standard that always applies to the normal distribution of IQ).

168. Gurin, 5 MICH. J. RACE & L. at 380.

significance, often an important metric, has become entirely divorced from *practical* significance.

The problem is certainly not confined to the CIRP study.[169] Most research regarding the educational benefits of diversity and affirmative action has used large, multi-institutional samples. Of course, once you have come close to testing *every* student at a given school, your descriptive generalizations about students at that school will be presumptively accurate. The confusion arises in generalizing *comparisons* between schools. When tests of large samples are reduced to pairwise comparisons, they risk being too powerful, unfairly stacking the cards to make small differences seem important.

G. *Experimenter Bias*

Part and parcel to self-report bias[170] is the possibility that an experimenter—again, consciously or unconsciously—may *elicit* particular responses from the participants in her study. To discuss experimenter bias is not necessarily to impeach an experimenter's motives or ethics. Still, a fundamental tenet of social science research is that an experimenter should have no personal stake in the outcomes; that she should be neutral, not an advocate. It should cause some alarm, then, that Gurin was the Psychology Department Chair at the University of Michigan—the *Grutter* and *Gratz* defendant—and she conducted the research specifically for those litigations. The Court, if it is to consider social science at all, should at least establish an admissibility threshold requiring that a study be performed by a researcher neutral to the litigation before the Court.[171]

H. *Required Time Limit for a Purportedly Permanent Educational Benefit*

Another gap in the diversity logic concerns Justice O'Connor's prophecy that "25 years from now, the use of racial preferences will no longer be necessary to further the interest approved today," a prediction

---

169. Even the MSS study of students solely from Michigan had a "sample" size of 1,321. *Id.*

170. *See supra*, Part IV.D.2.

171. Crystal Muhammad identifies some risks that self-interested research might pose:

> [Researchers] may be pressured by their personal ideology, concern for populations at risk, or stakeholders to generate desired results. The implication is that applied studies are generally not as valued as pure research. In addition, given the pressures of litigation in particular, courts tend to give greater credence to studies conducted outside of the context of litigation than projects ensued in the wake of litigation.

Muhammad, *supra* note 20, at 25 n.35 (citations omitted).

stemming from the requirement that "race-conscious admissions policies must be limited in time."[172] However, if diversity is theoretically linked to educational benefits, it should continue to be linked forever.[173] O'Connor apparently hopes that a critical mass of all races will soon be achieved by accident, without considering race. Such an accident would require that the Black-White standardized test score gap[174] disappear, that currently preferred minorities continue to apply in sufficient numbers to maintain their representation, and that different-race students become, on average, indistinguishable on paper. But then, race would no longer be relevant to education, contradicting the very foundation of the diversity rationale.[175] In the alternative, if diversity's benefit is durable, quota-like percentage sensitivity is inescapable, as is the acceptance that even Whites might become a preference group if campus demographics dictated the need.

I. *Diverse but Segregated*

The Education Brief defends the use of enrollment in an ethnic studies course as a diversity measure, arguing that, from overall membership in ethnic studies courses, one can abstract a university structure supportive of diversity.[176] Indeed, Gurin never argues that an institution's overall structural diversity should have a direct impact on education outcomes:

> Structural diversity is essential, but, by itself, usually not sufficient to produce substantial benefits; in addition to being together on the same campus, students from diverse backgrounds must also learn about each other in the courses they take and in informal interaction outside of the classroom. For

---

172. *Grutter*, 539 U.S. at 342–43.

173. *E.g.*, *Gratz*, 122 F. Supp. 2d at 824 ("[U]nlike the remedial setting, where the need for remedial action terminates once the effects of past discrimination have been eradicated, the need for diversity lives on perpetually."); Lauriat, *supra* note 107, at 1191 ("Unlike remedial measures addressing past discrimination, the diversity rationale would allow racial classifications to be used for an indefinite period of time, as long as an institution wished to assemble its student body along racial lines and dismisses race neutral alternatives after 'serious, good faith consideration.' " (quoting *Grutter*, 539 U.S. at 339)).

174. *See generally* CHRISTOPHER JENCKS & MEREDITH PHILLIPS, THE BLACK-WHITE TEST SCORE GAP (1998) (collection of chapters by authors who, in regard to White students' documented superiority over Black students' vocabulary, reading, and mathematics standardized test scores, offer a range of explanations for the score gap—from genetics to environment to test bias—and explore how and why the gap has narrowed over time).

175. An alternative end to affirmative action might come when "race is no longer a powerful influence on individual experiences (or is no longer sufficiently salient that its inclusion in education furthers learning)." *Evidentiary Framework*, *supra* note 120, at 1364.

176. Education Brief, *supra* note 68, at 9. Note how this argument confounds a supportive university structure with a causal hypothesis centered upon individual students.

> new learning to occur, institutions of higher education have to make appropriate use of structural diversity.[177]

But the compelling interest articulated in *Grutter* derives exclusively from an overall diverse student body. Now we are told that a diverse student body, by itself, is insufficient; that a university must *also* successfully manipulate that diversity, once numerically achieved, in order to realize any educational benefit. In short, Gurin's model requires a mediated relationship,[178] but the Court's reasoning does not. Thus, the diversity interest sustained by the Court contradicts the very evidence used to articulate the interest in the first place.

If the benefits of exposure to diverse viewpoints are real, then admitting a diverse class is not enough. The university should ensure that those diverse persons actually interact. A university's assertions that diversity promotes cross-cultural understanding, defeats stereotypes, and enriches one's education all suffer if the college provides, sponsors, or condones separate minority housing, student groups, orientation programs, or employer recruitment events.[179] Further, a university allowing students to liberally choose their courses will sacrifice some of diversity's supposed benefit whenever race correlates with course selection. At a university where, accidentally, all Black students major in humanities and all White students major in science, no Black student might ever share a class with a White student, yet the university could call itself "diverse." *Grutter* allows schools to stop trying once they have achieved mere quantitative diversity, and does not mandate "more meaningful, real equality that would require understanding, and rejecting dominance in favor of true egalitarian pluralism."[180]

I conclude this section the same way I began, with an observation by Paul Brest:

> These views [about the benefits of diversity] are based on my experience as a professor and dean at Stanford Law School . . . While they are supported by some empirical studies, *one would*

---

177. *Id.* at 10 (quoting Supplemental Expert Report of Patricia Y. Gurin, Grutter v. Bollinger, No. 97-75928 (E.D. Mich. 2001)).

178. *Id.* at 10 ("'[A] large body of social science research show[s] that institutional variables [such as structural diversity] have their effects on individual level variables [such as positive educational outcomes] through other mechanisms.'") (quoting EWART A.C. THOMAS & RICHARD J. SHAVELSON, ANALYSIS OF REPORT OF WOOD & SHERMAN, ADDENDUM TO NATIONAL ASSOCIATION OF SCHOLARS AMICUS BRIEF, *available at* http://siher.stanford.edu).

179. *See* Bunzel, *supra* note 3, at 493 (citing INSTITUTE FOR THE STUDY OF SOCIAL CHANGE, UNIVERSITY OF CALIFORNIA BERKELEY, DIVERSITY PROJECT (1991)); JOHN H. BUNZEL, RACE RELATIONS ON CAMPUS: STANFORD STUDENTS SPEAK (1992) (describing how Black Stanford University students in the early 1970s sought a haven from "white Stanford," and White students believed Blacks were "separating themselves in unhealthy ways").

180. Dowd et al., *supra* note 62, at 38.

> *have to concede that the evidence to date is largely impressionistic. . . .* [M]any educators believe that diversity is educationally valuable. But the evidence is impressionistic and the conclusions are speculative, or perhaps just hopeful. . . . The state of empirical knowledge about the educational benefits of diversity belies any claim of necessity.[181]

Brest's awareness (and humility) is rare on both sides of the impassioned debate over a continuing role for affirmative action in university admissions. He is able to separate his own theory about the value of classroom diversity from an honest assessment that the proven link between diversity and educational benefit is thin. It is possible, of course, to value diversity oneself, but still recognize that one's own values are not "compelling."

CONCLUSION

In *Grutter*, a Supreme Court majority authoritatively proclaimed what Justice Powell alone declared in *Bakke*: that racial diversity produces substantial educational benefits warranting racial preferences in public university admissions. In this Note I have shown how social science has not—indeed, *cannot*—prove that student body diversity produces these benefits. And I have argued that, where the diversity rationale cannot be proven, it likewise cannot be compelling.

It is possible the *Grutter* Court was using the diversity rationale as a subterfuge for retaining, in a qualified sense, affirmative action policies the Court finds desirable but has rejected on other grounds, including as a remedy for prior discrimination in society.[182] At best, such action is a useful fiction, sociologically supporting what everyone "knows" anecdotally to be just.[183] This approach did not seem dangerous before *Grutter*, when courts rarely relied on social science evidence in desegregation litigation.[184] However, when social science becomes central to a court's reasoning, as in *Grutter*, one might reasonably suspect that such evidence is

---

181. Brest, *supra* note 101, at 685, 690–91 (emphasis added) (citations omitted).

182. *See supra* notes 3–4, 11–12 and accompanying text.

183. *See, e.g.*, Charles Black, *The Lawfulness of the Segregation Decisions*, 69 YALE L.J. 421, 423–24, 426 (1960) (calling the lawfulness of the Supreme Court's segregation decisions "subjectively obvious", a "matter of common notoriety", and part of the "background knowledge of educated men"; responding with "laughter" to naysayers who would suggest that segregation does not offend equality; and arguing that Court decisions to the contrary would be "law based on self-induced blindness, on flagrant contradiction of known fact.").

184. Muhammad, *supra* note 20, at 2 (finding, in a review of 28 federal court desegregation decisions involving 38 judges, no significant relationship between case outcomes and the quality of social science evidence presented, and observing that social science only plays a role for cases and judges at the margin).

being used as a cover to lend an appearance of objectivity to a decision made on normative grounds, to dart political controversy. James Ryan has found, "[I]f judges perceive an issue as involving moral or philosophical judgments, as opposed to pragmatic or instrumental ones, they are less likely to rely heavily on social science evidence to resolve the issue."[185] The *Grutter* Court deviated from this approach.

Let us be honest: affirmative action is a fundamentally moral policy. If affirmative action is legal, it is legal because equal protection permits, maybe *demands*, that the gatekeepers of national power and wealth acknowledge and correct for their own past discriminatory actions, and that they grant future access to power and wealth to persons who might otherwise be excluded.[186] The diversity rationale, in contrast, is a purely functional justification, conspicuously lacking the moral component. It is a consolation prize.

"It is hard to be an enemy of diversity. Most Americans recognize diversity as one of the nation's proudest attributes."[187] I agree: diversity is a morally justified goal. Still, I disagree that diversity is an educationally justified means.[188] This latter claim remains only a "hypothesis in search of proof."[189] In summary, then, this Note can fairly be characterized as a response to the following:

> The benefits of diversity are the result of interpersonal interactions that cannot be quantified or verified by scientific proof.

---

185. Ryan, *supra* note 27, at 1662–63.

186. *See generally* Lawrence III, *Two Views of the River*, *supra* note 4 (lamenting the amoral aspect of the diversity rationale, in contrast to more radical arguments focusing on the need to remedy past discrimination or address present discriminatory practices; and arguing that the diversity rationale, rather than counteract traditional notions of merit that serve to perpetuate race and class privilege, actually facilitates the reproduction of elites).

187. Bunzel, *supra* note 3, at 498.

188. This distinction was understood by scholars attending conferences held by the Harvard Civil Rights Project following the *Hopwood* decision, where the lack of solid evidence to promote the case for diversity was addressed: "Legal experts poked . . . holes in other pieces of research, disheartening some of the academics in attendance, who were confronted with the need to justify a concept they believe in implicitly." Douglas Lederman, *Backers of Affirmative Action Seek Research To Bolster Cause: At Meeting at Harvard, Sympathetic Lawyers Tell Scholars That Their Work Will not Sway Many Judges*, CHRON. OF HIGHER EDUC., May 23, 1997, at A28.

189. Bunzel, *supra* note 3, at 503. Bunzel continues:

> I am prepared, therefore, to entertain two thoughts simultaneously: 1) admitting students from different geographical regions, races, and social backgrounds is a desirable goal and a university's public responsibility, but 2) diversity of viewpoints in the classroom comes principally from intellectual effort on the part of students and professors working together to explore ideas, test arguments, and question assumptions, with one's race or ethnicity rarely the determinative factor.

Id., at 505.

> Such benefits do not accrue every day to every student and may even go undetected by the students and participants involved. Therefore, a university would face insurmountable obstacles if it had to 'prove' the benefits of diversity using current methodologies.... Rather than preventing universities from seeking the educational benefits of diversity, courts should find the testimony of educators sufficient to establish these benefits.... [A] university should not be required to make a particularized showing that diversity furthers learning at its institution.[190]

I am unwilling to excuse a lack of evidence and simply trust the word of educators. A Court that accepts ignorance and faith as a substitute for principle and proof is irresponsible. After all, "[W]e must never forget, that it is a constitution we are expounding."[191]

190. *Evidentiary Framework, supra* note 120, at 1361–62 (citations omitted). For critique of this argument, *see* Thomas E. Wood & Malcolm J. Sherman, *Is Campus Racial Diversity Correlated with Educational Benefits?*, *in* NATIONAL ASSOCIATION OF SCHOLARS, RACE AND HIGHER EDUCATION: WHY JUSTICE POWELL'S DIVERSITY RATIONALE FOR RACIAL PREFERENCES IN HIGHER EDUCATION MUST BE REJECTED (2001), *available at* http://www.nas.org/rhe.pdf.

191. McCulloch v. Maryland, 17 U.S. 316, 407 (1819).