

Boston University School of Law

Scholarly Commons at Boston University School of Law

Faculty Scholarship

Fall 2011

Racing Towards Colorblindness: Stereotype Threat and the Myth of Meritocracy

Jonathan P. Feingold

Boston University School of Law

Follow this and additional works at: https://scholarship.law.bu.edu/faculty_scholarship



Part of the [Law and Race Commons](#), and the [Legal Education Commons](#)

Recommended Citation

Jonathan P. Feingold, *Racing Towards Colorblindness: Stereotype Threat and the Myth of Meritocracy*, 3 *Georgetown Journal of Law and Modern Critical Race Perspectives* 231 (2011).

Available at: https://scholarship.law.bu.edu/faculty_scholarship/826

This Article is brought to you for free and open access by Scholarly Commons at Boston University School of Law. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarly Commons at Boston University School of Law. For more information, please contact lawlessa@bu.edu.



NOTES

Racing Towards Color-blindness: Stereotype Threat and the Myth of Meritocracy

JONATHAN FEINGOLD*

INTRODUCTION

Stereotype threat may be a possible source of bias in standardized tests, a bias that arises not from item content but from group differences in the threat that societal stereotypes attach to test performance.

—Claude Steele¹

Imagine a faulty Scantron machine.² The defective machine adds four points to the Law School Admissions Test (LSAT) scores of White students. What if a law school's admissions office knowingly used this machine? Such behavior would be morally and constitutionally suspect. While the school is not intentionally discriminating on the basis of race, it is relying on a defective tool that precludes a race-neutral, individualized review of every applicant. Aware that the United States Supreme Court recently invalidated a similar policy,³ the school decides to fix their defective admissions process. However, the school is unable to stop the Scantron from allocating points to White students. Fortunately, it is able to adjust the machine so that it adds four points to the scores of all non-White students. Now, all White and non-White students receive a four-point bump.

Could the Constitution forbid this type of race-consciousness? Answering this question begins by turning to a largely forgotten component of Justice Powell's opinion in *Regents of the University of California v. Bakke*.⁴ In *Bakke*, Justice Powell provided the decisive vote in the Court's invalidation of the UC Davis Medical School's race-conscious admissions program.⁵ Beyond his holding, the legal community most remembers *Bakke* for Justice Powell's articulation of student body diversity as a compelling state interest.⁶ However, this emphasis on the diversity rationale has obscured a companion compelling interest identified by Justice Powell. In addition

* J.D. Candidate, UCLA School of Law 2012. I thank Jerry Kang for his guidance and support throughout the writing of this piece. Devon Carbado, Nilanjana Dasgupta, Rachel Godsil, Cheryl Harris, Luke Harris, Kimberle Crenshaw, and Karen Lorang provided invaluable comments on earlier drafts. © 2012, Jonathan Feingold.

1. Claude M. Steele, *A Threat in the Air*, 6 AM. PSYCHOL. 613, 622 (1997).
2. A Scantron is the machine that calculates standardized test scores.
3. See *Gratz v. Bollinger*, 539 U.S. 244 (2003) (invalidating the University of Michigan's undergraduate admissions program for awarding twenty points to underrepresented minorities on the basis of race).
4. *Regents of Univ. of Cal. v. Bakke*, 438 U.S. 265 (1978).
5. *Id.* at 314.
6. *Id.* at 311-15 (holding that student body diversity in higher education is a compelling government interest).

to student body diversity, Justice Powell suggested that “[r]acial classifications in admissions could serve a fifth purpose . . . fair appraisal of each individual’s academic promise in the light of some cultural bias in grading or testing procedures.”⁷

This statement reveals Justice Powell’s understanding that measurement biases have the potential to corrupt the tools, such as admissions tests, that we view as neutral and meritocratic. Thus, as if imagining the faulty Scantron machine, Justice Powell concluded that if “race and ethnic background were considered only to the extent of curing established inaccuracies in predicting academic performance, it might be argued that [race-conscious admissions constitute] no ‘preference’ at all.”⁸ Justice Powell recognized that when beginning from a baseline of defective measures, race-conscious interventions that correct the defect actually mitigate preference and promote merit. This understanding of race-consciousness presents an important opportunity for self-reflection from color-blind jurisprudence.⁹ If color-blindness dictates that public institutions of higher education should never make decisions on the basis of race, what does color-blindness prescribe if we discover that the current law school admissions regime relies on a tool as defective as our hypothetical Scantron machine?¹⁰

While there may be a debate over the particulars, most agree that the general concepts of fairness and merit should drive law school admissions. I am not entering the fairness debate. This Note is focused solely on merit.¹¹ For the purposes of my argument, I accept the normative claim that the allocation of scarce resources (e.g. admission to elite law schools) should occur on the basis of merit as measured by traditional indicia such as the LSAT.¹² However, this concession to meritocracy demands that we rely on “fair measures.”¹³ At a basic level, “fair measures” exist

7. *Bakke*, 438 U.S. at 306 n.43 (1978). To support my claim that the legal world has “forgotten” about this component of Justice Powell’s opinion, I point to the dearth of academic scholarship engaging this “test bias” rationale. A Westlaw search revealed 1,615 articles mentioning the diversity rationale [search terms: diversity & “compelling interest” & “438 U.S. 265”], yet only ten articles mentioning “test bias” rationale [search terms: “to the extent that race and ethnic background”].

8. *Id.*

9. “Color-blindness” rejects governmental recognition of race. Chief Justice Roberts exemplified this notion in *Parents Involved*. See *Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 748 (2007) (“The way to stop discriminating on the basis of race is to stop discriminating on the basis of race.”). For a critique of color-blindness see Laurence Tribe, *In What Vision of the Constitution Must the Law be Color-Blind*, 20 J. MARSHALL L. REV. 201 (1986).

10. While it is unclear exactly what the LSAT measures, I employ “talent” and “merit” throughout this Note to refer to the highly valued attributes that the LSAT supposedly measures. Since my fair-measures argument demands only that the LSAT equally and accurately measure the “talent” of all individuals, it is unnecessary to define this term with greater precision.

11. A critique of our common conceptions of merit is unnecessary for the purposes of this Note. I thus avoid that discussion. Still, it is important to recognize that merit is a suspect concept that has acquired the veneer of neutrality and objectivity through a process of normalization and naturalization. See, e.g., Daria Roithmayr, *Deconstructing the Distinction Between Bias and Merit*, 85 CALIF. L. REV. 1449 (1997).

12. I am making an argument-specific concession that LSAT scores are an acceptable proxy for legal talent, one factor that should be considered when making admissions decisions.

13. Jerry Kang & Mahzarin Banaji, *Fair Measures: A Behavioral Realist Revision of Affirmative Action*, 94 CALIF. L. REV. 1063, 1067 (2006) (“Fair” connotes the moral intuition that being fair involves an absence of unwarranted discrimination, by which we mean unjustified social category-contingent behavior. The term

when our tools accurately measure the talent of every student without regard to social categories such as race.

Recent social-science findings suggest that the LSAT is not meeting this demand.¹⁴ This failure of fair measures is due to stereotype threat,¹⁵ a psychological phenomenon that depresses the performance of negatively stereotyped groups when performing stereotype-relevant tasks.¹⁶ When exposed to stereotype threat, vulnerable Black and Latino/a LSAT-takers suffer from a race-dependent measurement bias.¹⁷ Functionally, this measurement bias is no different than the faulty Scantron machine discussed above. In the presence of stereotype threat, the LSAT is a defective tool that prevents otherwise qualified Black and Latino/a students from displaying their true talent. Understood in this way, reliance on the LSAT contravenes fair measures and deprives Black and Latino/a applicants of the individualized review our Constitution demands.¹⁸

The harm arising from this “fair measures” defect extends beyond those students who would have gained admission but for stereotype threat. As a consequence of the disproportionate exclusion of talented Black and Latino/a students, the few Blacks and Latinos/as who access elite institutions face unique burdens associated with their token status.¹⁹ A holistic understanding of the harm produced by reliance on a defective measure of talent requires broadening our view beyond the single moment of admissions decision-making.

Part I of this Note explores the psychological phenomenon of stereotype threat and its implications regarding the performance of Black and Latino/a students on the LSAT and in law school. Taking the science seriously complicates our historical acceptance of the LSAT as a neutral judge of talent. Recognizing these issues, Part II presents rescaling, an intervention that cures stereotype threat-induced measurement biases. Specifically, rescaling offsets the mismeasurement of vulnerable Black and Latino/a students by correcting scores in accordance with the observed mean effect of stereotype threat on LSAT-takers. As this measure entails express race-consciousness, policy and legal objections will inevitably arise. Responding first to policy concerns, I argue that rescaling is not solely a “fair measures” remedy. Rescaling additionally enables the attainment of a critical mass of underrepresented students. Critical mass prevents the manifestation of racially hostile environments in law school classrooms that carry their own set of race-dependent burdens. Thus, by relying on a more accurate measure of Black and Latino/a talent at the admissions stage, rescaling also

also connotes accuracy in assessment. ‘Measure’ has the double meaning as well: measurement and an intervention intentionally taken to solve a problem.”).

14. This claim is not limited to the LSAT. Stereotype threat is likely to affect vulnerable individuals across all highly time-pressured standardized tests.

15. For an accessible review of stereotype threat see CLAUDE STEELE, *WHISTLING VIVALDI: AND OTHER CLUES TO HOW STEREOTYPES AFFECT US* (2010).

16. See *infra* Part I.

17. In other words, an LSAT score of 165 for a vulnerable, stereotype-threatened Black student is not the same as a score of 165 for their White peer. See STEELE, *supra* note 15.

18. See *Grutter v. Bollinger*, 539 U.S. 306, 334-37 (2003); *Gratz v. Bollinger*, 539 U.S. 244, 268-76 (2003); *Regents of Univ. of Calif. v. Bakke*, 438 U.S. 265, 315-16 (1978).

19. In other words, fair measure defects inhibit the attainment of a diverse student body.

facilitates the attainment of diverse student bodies that foster a more race-neutral and level playing field within legal institutions.

Having connected the complementary and interdependent goals of fair measures and critical mass, Part III presents a constitutional defense of rescaling. I argue that our knowledge about stereotype threat breathes new life into Justice Powell's forgotten insight from *Bakke*. Such resuscitation does not require creating a "new" compelling interest. However, the reality of stereotype threat and its relation to underrepresentation in law school calls for an updated understanding of the well-established compelling interests of diversity and remedying discrimination. Our new understanding of stereotype threat contests the notion that critical mass (i.e. the achievement of meaningful diversity) and fair measures (i.e. the eradication of discrimination) function independently of one another. Rather, since the promotion of fair measures enables the achievement of a critical mass, I present a fair measures-critical mass frame that recognizes the interdependence of these intimately connected goals. Understood through this hybrid diversity-remedying discrimination lens, rescaling emerges as a constitutional policy that directly promotes these historically recognized interests.

I. STEREOTYPE THREAT

Like many standardized tests, the LSAT continues to produce significant racial disparities.²⁰ In the 2008-2009 school year, White LSAT-takers scored on average ten points higher than Blacks and six points higher than Latino/as (SD = 8.94).²¹ Various rationales have been offered to explain the racial gap in standardized testing: genetic inferiority,²² lack of preparation,²³ lack of effort, inferior schooling,²⁴ cultural deficiency,²⁵ and family structure and poverty.²⁶ These explanations, whether implicating nature or nurture, share the common assumption that lower scoring minorities lack merit, and thus, do not belong in elite institutions.²⁷

20. See LSAT TECHNICAL REPORT SERIES, LSAT PERFORMANCE WITH REGIONAL, GENDER, AND RACIAL/ETHNIC BREAKDOWNS: 2003-2004 THROUGH 2009-2010 TESTING YEARS, <http://www.lsac.org/LsacResources/Research/TR/TR-10-03.pdf>.

21. *Id.* at 19, tbl. 4A.

22. See, e.g., RICHARD HERRNSTEIN & CHARLES MURRAY, *THE BELL CURVE* (1994); ARTHUR JENSEN, *EDUCABILITY AND GROUP DIFFERENCES* (1973).

23. See STEPHAN THERNSTROM & ABIGAIL THERNSTROM, *AMERICA IN BLACK AND WHITE: ONE NATION, INDIVISIBLE* 402-03 (1997).

24. See, e.g., Ronald G. Fryer Jr. & Steven D. Levitt, *Understanding the Black-White Test Score Gap in the First Two Years of School*, 86 REV. ECON. & STAT. 447 (2004).

25. See, e.g., Philip J. Cook & Jens Ludwig, *The Burden of Acting White: Do Black Adolescents Disparage Academic Achievement?* in *THE BLACK-WHITE TEST SCORE GAP* 375 (Christopher Jencks & Meredith Phillips eds., 1998); Signthia Fordham & John Ogbu, *Black Students' School Successes: Coping with the Burden of Acting White*, 18 URBAN REV. 176 (1986).

26. See CONSEQUENCES OF GROWING UP POOR (Greg J. Duncan & Jeanne Brooks-Gunn, eds., 1997); SUSAN MAYER, *WHAT MONEY CAN'T BUY: FAMILY INCOME AND CHILDREN'S LIFE CHANCES* (1997); Meredith Phillips et al., *Family Background, Parenting Practices, and the Black-White Test Score Gap*, in *THE BLACK-WHITE TEST SCORE GAP* 375 (Christopher Jencks & Meredith Phillips eds., 1998).

27. See STEREOTYPE THREAT: THEORY, PROCESS, AND APPLICATION (Michael Inzlicht & Toni Schmader, eds., 2011).

Even if socioeconomic realities and unequal access to resources produce part of the achievement gap, these explanations prove incomplete. A gap remains even when controlling for the above factors.²⁸ Race plays a role. Contrary to biological explanations, however, underperformance is not the result of the inherent inferiority of Black and Latino/a students.²⁹ Rather, this otherwise unexplained portion of the racial gap is, in part, a consequence of stereotype threat, a psychological phenomenon tied to social stereotypes and the racial identities of Black and Latino/a students.³⁰

Stereotype threat is defined as the “social-psychological threat that arises when one is in a situation or doing something for which a negative stereotype about one’s group applies.”³¹ In their book on stereotype threat, Professors Michael Inzlicht and Toni Schmader explain that years of research has shown that the “fear of stereotype confirmation can hijack the cognitive systems required for optimal performance, resulting in low test performance.”³²

A. *The Studies*³³

African Americans and Intellectual Ability. In their seminal study, Claude Steele and Joshua Aronson focused on the interaction between racial identity, stereotype, and the cognitive performance of highly talented and confident Black students from Stanford University.³⁴ Steele and Aronson administered thirty-minute tests composed of verbal GRE questions to White and Black participants.³⁵ They varied the

28. See Fryer Jr. & Levitt, *supra* note 24, at 447 (“Even after controlling for a wide range of covariates including family structure, socioeconomic status, measures of school quality, and neighborhood characteristics, a substantial racial gap in test scores persists.”); Claude M. Steele, *Race and the Schooling of Black Americans*, 269 ATLANTIC MONTHLY 68 (1992) (arguing that standard explanations are incomplete, as low SAT scorers are not more likely to flunk out than high SAT scores; achievement gaps exist even when Blacks do not suffer financial disadvantage; and controlling for any level of previous school preparation, Blacks achieve less in subsequent schooling than do Whites).

29. See Kimberly West-Faulcon, *The River Runs Dry: When Title VI Trumps State Affirmative Action Laws*, 157 U. PA. L. REV. 1075, 1109-17 (2009) (demonstrating that when socioeconomic factors are taken into account yet discrepancies remain, minority deficiency-based explanations inherently rely on biological notions of race); Devon M. Carbado, *Yellow by Law*, 97 CALIF. L. REV. 633 (2009) (exposing the fragility of biological arguments).

30. See Gregory M. Walton & Steven J. Spencer, *Latent Ability: Grades and Test Scores Systematically Underestimate the Intellectual Ability of Negatively Stereotyped Students*, 20 PSYCHOL. SCI. 1132, 1132 (2009) (“[A]t least a portion of group differences is illusory . . . this results from pervasive psychological threats in academic environments, which undermine the performances of ethnic minority students and of women.”). Walton and Spencer observed an effect size for Blacks and Latino/as on the Math and Reading SAT ranged between 39 and 41 points, with an overall Black/White and Latino/a/White disparity of 199 and 148 points respectively. *Id.* at 1137.

31. Steele, *supra* note 1, at 614.

32. Inzlicht & Schmader, *supra* note 27.

33. There are well over one hundred published studies on stereotype threat. I chose to highlight these particular studies primarily because they successfully illustrate the contextual nature of stereotype threat. They also show stereotype threat is not limited to racial or historically subordinated groups and that stereotype threat is most burdensome on the vanguard of a group.

34. See Claude M. Steele & Joshua Aronson, *Stereotype Threat and the Intellectual Test Performance of African Americans*, 69 J. PERSONALITY & SOC. PSYCHOL. 797 (1995).

35. *Id.* at 799.

level of stereotype relevance by presenting the task in multiple ways. Half of the students were told the activity measured intellectual aptitude. Priming students in this way increased the relevance of negative stereotypes about Black cognitive ability. For the other half, Steele and Aronson negated the relevance of societal stereotypes by presenting the task as a problem-solving experiment unrelated to innate ability.³⁶ The results suggested that this slight environmental change—presenting an activity as diagnostic or non-diagnostic of a negatively stereotyped trait—was sufficient to affect performance. Black students in the “intellectual ability” group performed significantly worse than their White counterparts.³⁷ However, Black students in the “problem solving” group performed at the level of their White counterparts.³⁸

(Asian) Women and Math. A gender achievement gap persists in the fields of mathematics and the hard sciences. Traditional explanations invoke notions of inherent gender differences³⁹ and gender-role socialization.⁴⁰ Similar to the rationales explaining the racial achievement gap, these explanations fail to tell the entire story.

A pair of 1999 studies began to fill in the gap.⁴¹ In the first study, researchers selected a group of men and women undergraduates highly skilled in math.⁴² The participants completed either a very difficult or a relatively easy math test.⁴³ Women performed significantly worse than men on the difficult test.⁴⁴ However, the easier test produced no gender difference.⁴⁵ Did the results suggest a biological ceiling on women’s math ability? Or, had something extraneous to innate talent inhibited the women’s performance on the more difficult exam?

To answer this question, the researchers re-administered the difficult test to all participants. This time, however, they varied the degree to which gender stereotypes were relevant to performance.⁴⁶ For the first group, administrators primed a negative stereotype about women by informing participants that the test had previously shown gender differences.⁴⁷ For the second group, administrators removed any relevance of math-gender stereotypes by explaining that the test had never shown gender differ-

36. *Id.*

37. *Id.* at 801 (controlling for SAT).

38. *Id.* at 801-02.

39. See, e.g., Lawrence Summers, Sciences Remarks at NBER Conference on Diversifying the Science & Engineering Workforce (Jan. 14, 2005), <http://web.archive.org/web/20080130023006/http://www.president.harvard.edu/speeches/2005/nber.html>.

40. See, e.g., Janet K. Swim, *Perceived Versus Meta-Analytic Effect Sizes: An Assessment of the Accuracy of Gender Stereotypes*, 66 J. PERSONALITY & SOC. PSYCHOL. 21 (1994); Jacquelynne S. Eccles, Janis E. Jacobs & Rena D. Harold, *Gender Role Stereotypes, Expectancy Effects, and Parents’ Socialization of Gender Differences*, 46 J. SOC. ISSUES 183 (1990).

41. See Steven J. Spencer, Claude M. Steele & Diane M. Quinn, *Stereotype Threat and Women’s Math Performance*, 35 J. EXPERIMENTAL SOC. PSYCHOL. 4, 7 (1999); Margaret Shih, Todd L. Pittinsky & Nalini Ambady, *Stereotype Susceptibility: Identity Salience and Shifts in Quantitative Performance*, 10 PSYCHOL. SCI. 80 (1999).

42. Spencer, Steele & Quinn, *supra* note 42, at 8 (selecting students who scored above the 85th percentile on the SAT Math section and strongly identified with math).

43. *Id.*

44. *Id.* at 10.

45. *Id.*

46. *Id.*

47. *Id.* at 10-11.

ences.⁴⁸ This seemingly minor instruction greatly impacted results. While the “gender difference” group retained a significant disparity,⁴⁹ women in the “no-difference” group performed at the same level as men.⁵⁰ Thus, innate ability was not responsible for the previously observed gender gap. Underperformance was not the result of biology or a lack of ability.

If membership in a negatively stereotyped group undermines performance, could membership in a positively stereotyped group boost performance? To answer this question, researchers examined the interplay between the conflicting identities of Asian American women in the domain of math.⁵¹

Asian American female undergraduates answered a set of difficult math questions.⁵² Prior to each test, the students filled out one of three questionnaires. The questionnaires were intentionally designed to either: 1) prime gender identity;⁵³ 2) prime racial identity;⁵⁴ or 3) prime neither identity.⁵⁵ The results confirmed that the relevant stereotype, as opposed to innate ability, dramatically predicted performance. The control group answered an average of 49% of questions correctly, whereas the race-prime group answered 54% correctly and the gender-prime group answered only 43% correctly.⁵⁶ To verify the perceived role of stereotype threat, researchers recreated the experiment in Vancouver, Canada, a community where the “Asians are good at math” stereotype does not exist.⁵⁷ Unlike the American study, the race-prime group experienced a depression in performance.⁵⁸ These findings added support for the theory that the specific societal stereotype, not race itself, was the main culprit affecting performance.

White Men and Math. The preceding studies illustrate stereotype threat’s ability to affect the performance of individuals from historically stigmatized groups. But it is not just minorities and women who are vulnerable to stereotype threat.

A set of studies revealed this equal-opportunity characteristic of stereotype threat by observing its effect on highly math-proficient White male undergraduates.⁵⁹ The

48. *Id.* at 11.

49. *Id.* at 13.

50. *Id.*

51. See Shih, Pittinsky & Ambady, *supra* note 41, at 80. “Conflicting identities” refers to the positive stereotype associated with Asians and math and the negative stereotype associated with women and math.

52. *Id.* at 81 (selecting students with a mean Quantitative SAT score of 750.9).

53. *Id.* (asking gender-specific questions).

54. *Id.* (asking questions about language and family genealogy).

55. *Id.* (asking questions about phone and cable service).

56. *Id.* All participants demonstrated similar levels of motivation, thus supporting the notion that underperformance was not caused by a lack of effort.

57. *Id.* at 82. This stereotype does not exist because Vancouver’s Asian community is composed largely of recent immigrants. *Id.*

58. *Id.* (observing that while the control group answered 59% of the questions accurately, both the racial-prime and gender-prime groups answered correctly only 44% and 29% respectively).

59. See Joshua Aronson et al., *When White Men Can’t Do Math: Necessary and Sufficient Factors in Stereotype Threat*, 35 J. EXPERIMENTAL SOC. PSYCHOL. 29, 33 (1999) (selecting students with an average Math SAT score of 712 and who strongly identified with the domain of math). See also, Jessi L. Smith & Paul H. White, *An Examination of Implicitly Activated, Explicitly Activated, and Nullified Stereotypes on Mathematical Performance: It’s Not Just a Woman’s Issue*, 47 SEX ROLES 179 (2002).

participants completed a set of difficult math questions.⁶⁰ Mirroring earlier studies, administrators varied the relevance of identity through distinct presentations of the task. For the first group, administrators characterized the study as an attempt to understand Asian superiority in math.⁶¹ For the second group, administrators did not mention Asians. They stated solely that the test was a measure of math ability.⁶² Consistent with previous findings, the mere situational change (addition of Asian-positive stereotype) was enough to disrupt performance. Participants in the first group answered an average of 36% questions correctly, while participants in the second group answered over 53% correctly.⁶³

B. "A Threat in the Air"

The foregoing examples comprise only a small sample of the many studies documenting the adverse effects of stereotype threat.⁶⁴ This research has documented stereotype threat across numerous axes of identity: including socioeconomic status,⁶⁵ age,⁶⁶ gender,⁶⁷ and race.⁶⁸ Each study tells a similar story. When individuals con-

60. Aronson et al., *supra* note 59, at 34 (coming from the GRE mathematics subject test).

61. *Id.* at 33 (providing participants with articles from noteworthy sources that emphasized the growing math achievement gap between Asians and Whites).

62. *Id.* at 34.

63. *Id.*

64. For two comprehensive meta-analyses, see Hannah-Hanh D. Nguyen & Ann M. Ryan, *Does Stereotype Threat Affect Test Performance of Minorities and Women? A Meta-Analysis of Experimental Evidence*, 93 J. APPLIED PSYCHOL. 1314 (2008) (analyzing 116 studies from 76 published reports); Walton & Spencer, *supra* note 30 (analyzing thirty-nine studies on stereotype threat). Thirty years before Steele's seminal work, studies had uncovered findings that we can now understand as the product of stereotype threat. For instance, Katz, Roberts, and Robinson found that Black students performed better on an IQ test when presented as diagnostic of hand-eye coordination than when presented as diagnostic of intelligence. See Irwin Katz, S. Oliver Roberts & James M. Robinson, *Effects of Task Difficulty, Race of Administrator, and Instructions on Digit-Symbol Performance of Negroes*, 2 J. PERSONALITY & SOC. PSYCHOL. 53 (1965).

65. See, e.g., Jean-Claude Croizet & Theresa Claire, *Extending the Concept of Stereotype Threat to Social Class: The Intellectual Underperformance of Students from Low Socioeconomic Backgrounds*, 24 PERSONALITY & SOC. PSYCHOL. BULL. 588 (1998) (observing the effect of stereotype threat in low socio-economic status French students); Bettina Spencer & Emanuele Castano, *Social Class is Dead. Long Live Social Class! Stereotype Threat Among Low Socioeconomic Status Individuals*, 20 SOC. JUST. RES. 418 (2007) (observing performance decrements for low socio-economic status individuals when negative stereotypes are made relevant).

66. See, e.g., Becca Levy, *Improving Memory in Old Age Through Implicit Self-Stereotyping*, 71 J. PERSONALITY & SOC. PSYCHOL. 1092 (1996) (observing that the implicit activation of positive stereotypes concerning aging improved memory performance in elderly participants, while an implicit activation of negative stereotypes worsened memory performance); Thomas M. Hess et al., *The Impact of Stereotype Threat on Age Differences in Memory Performance*, 57 J. GERONTOLOGY: PSYCHOL. SCI. P3 (2002) (observing that contextual factors affected memory performance).

67. See, e.g., Linette M. McJunkin, *Effects of Stereotype Threat on Undergraduate Women's Math Performance: Participant Pool vs. Classroom Situations*, 45 EMPORIA STATE RES. STUD. 27 (2009) (observing that female students experienced a significant improvement on math tests in classroom and laboratory settings when the relevance of negative stereotypes were explicitly alleviated); Toni Schmader, *Gender Identification Moderates Stereotype Threat Effects on Women's Math Performance*, 38 J. EXPERIMENTAL SOC. PSYCHOL. 194 (2001) (observing that domain identification mediated stereotype threat effects on women); Johannes Keller & Dirk Dauenheimer, *Stereotype Threat in the Classroom: Dejection Mediates the Disrupting Threat Effect on Women's Math Performance*, 29 PERSONALITY & SOC. PSYCHOL. BULL. 371 (2003) (observing sixth grade girls underperformed on math tests when "reminded" that girls were bad at math).

68. See, e.g., Joshua Aronson, Carrie B. Fried & Catherine Good, *Reducing the Effects of Stereotype Threat*

front tasks capable of confirming a negative group stereotype, performance suffers. However, when subtle environmental manipulations remove this threat, the performance decrement disappears. This situational nature of stereotype threat led Professor Claude Steele to label the phenomenon a “threat in the air.”⁶⁹ Identifying the debilitating effect of situational racial contingencies disrupts traditional explanations for the achievement gap on standardized tests. Under such conditions, tests like the LSAT are not neutral evaluators of talent and underperformance is not the result of inherent minority deficiencies.

The unique burdens encountered by stereotype-threatened individuals produce measureable physiological responses.⁷⁰ The particular mechanism that causes stereotype threat has been described as a “disruptive mental load”⁷¹ that inhibits “working memory.”⁷² Working memory refers to the cognitive processes that enable an individual to “control the focus of one’s attention and regulate behavior.”⁷³ While almost all individuals experience anxiety and stress on high stakes tests, this particular disruption of working memory uniquely burdens stereotype-threatened individuals. Steele

on African American College Students by Shaping Theories of Intelligence, 38 J. EXPERIMENTAL SOC. PSYCH. 113 (2002); Jeff Stone et al., *Stereotype Threat Effects on Black and White Athletic Performance*, 77 J. PERSONALITY & SOC. PSYCHOL. 1213 (1999) (“Black participants performed significantly worse than did control participants when performance on a golf task was framed as diagnostic of ‘sports intelligence.’ In comparison, White participants performed worse than did control participants when the golf task was framed as diagnostic of ‘natural athletic ability.’”); Jeff Stone, *Battling Doubt by Avoiding Practice: The Effects of Stereotype Threat on Self-Handicapping in White Athletes*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 1667 (2002); Patricia M. Gonzalez, Hart Blanton & Kevin J. Williams, *The Effects of Stereotype Threat and Double-Minority Status on the Test Performance of Latino Women*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 659 (2002) (demonstrating negative effects of stereotype threat on Latina participants).

69. Steele, *supra* note 1. See also, STEELE, *supra* note 15, at 93 (“[W]hatever the skills or vulnerability a group may have, situational differences in stereotype threat alone—contingencies of social identity—are fully sufficient to affect intellectual performance substantially.”).

70. See, e.g., STEELE, *supra* note 15, at 121 (“[P]art of stereotype threat’s effect . . . is caused directly by its effect of increasing heart rate, blood pressure, and related physiological signs of anxiety to the point that these reactions interfere with performance.”); Jim Blascovich et al., *African Americans and High Blood Pressure: The Role of Stereotype Threat*, 12 PSYCHOL. SCI. 225 (2001) (observing a correlation between higher blood pressure of stereotype-threatened Blacks and worse performance on difficult test items as compared to Whites and non-stereotype-threatened Blacks); Wendy Berry Mendes et al., *Challenge and Threat During Social Interactions with White and Black Men*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 939 (2002) (observing that White students forced to speak with an unfamiliar Black student exhibited a substantial increase in blood pressure as compared with White students meeting an unfamiliar White student).

71. See Jean-Claude Croizet et al., *Stereotype Threat Undermines Intellectual Performance by Triggering a Disruptive Mental Load*, 30 PERSONALITY & SOC. PSYCHOL. BULL. 721 (2004).

72. See Toni Schmader, Michael Johns & Chad Forbes, *An Integrated Process Model of Stereotype Threat Effects on Performance*, 115 PSYCHOL. REV. 336 (2008); Toni Schmader & Michael Johns, *Converging Evidence that Stereotype Threat Reduces Working Memory Capacity*, 85 J. PERSONALITY & SOC. PSYCHOL. 440 (2003). Stereotype-threatened women displayed lessened activity in the areas of the brain associated with mathematical learning, yet increased activity in areas associated with social and emotional processing. See Anne C. Krendl et al., *The Negative Consequences of Threat: A Functional Magnetic Resonance Imaging Investigation of the Neural Mechanisms Underlying Women’s Underperformance in Math*, 19 PSYCHOL. SCI. 168 (2008). Steele comments that this “increase[d] vigilance toward possible threat and bad consequences in the social environment, [thus] diverts attention and mental capacity away from the task at hand, which worsens performance and general functioning, all of which further exacerbates anxiety, which further intensifies the vigilance of threat and the diversion of attention.” STEELE, *supra* note 15, at 123-26.

73. Schmader, Johns & Forbes, *supra* note 72, at 340.

referred to this additional burden, which is limited to individuals facing the threat of confirming a stereotype, as a “pioneer tax.”⁷⁴ Working memory is especially relevant to standardized tests because this particular cognitive process becomes essential on speeded tasks that push individuals to the limits of their ability.⁷⁵ Highly speeded and difficult tests such as the LSAT, law school finals, and the Bar exam are precisely the types of situations where we would expect stereotype threat to be most disruptive.⁷⁶

Two elements of stereotype threat warrant emphasis. First, while we intuitively believe that better, smarter, and more perseverant students will be better able to overcome obstacles such as stereotype threat, the opposite is true.⁷⁷ This is the underlying irony of stereotype threat. As a result of their high “domain identification,”⁷⁸ the confident and talented students we would expect to overcome psychological hurdles are actually the most susceptible.⁷⁹ For the vanguard, the desire and commitment to outperform or disprove stereotypes often produces “over-efforting” that counterproductively distracts attention from the particular task at hand.⁸⁰ Thus, while we may be inclined to tell students to try harder, such advice may only exacerbate the problem.

74. Claude M. Steele, Expert Report of Claude M. Steele, (Jan. 15, 2011), available at <http://www.vpcomm.umich.edu/admissions/legal/expert/steele.html> (“[Stereotype-threatened students] pay an extra tax on their investment there, a ‘pioneer tax,’ if you will, of worry and vigilance that their futures will be compromised by the ways society perceives and treats their group. And it is paid everyday, in every stereotype-relevant situation.”).

75. See Schmader, Johns & Forbes, *supra* note 73, at 340-41. See also Laurie T. O’Brien & Christian S. Crandall, *Stereotype Threat and Arousal: Effects on Women’s Math Performance*, 29 PERSONALITY & SOC. PSYCHOL. BULL. 782 (2003) (observing that stereotype threat adversely affected women’s performance on a difficult math test, yet women performed better under the same conditions when taking an easy math test); Talia Ben-Zeev, Steven Fein & Michael Inzlicht, *Arousal and Stereotype Threat*, 41 J. EXPERIMENTAL SOC. PSYCHOL. 174 (2005) (extending and confirming the O’Brien & Crandall study).

76. See John L. Horn & Nayena Blankson, *Foundations for Better Understandings of Cognitive Abilities*, in CONTEMPORARY INTELLECTUAL ASSESSMENT: THEORIES, TESTS AND ISSUES 41 (Dawn P. Flanagan & Patti L. Harrison, eds., 2d ed. 2005).

77. See Steele, *supra* note 1, at 617 (“Stereotype threat affects only a subportion of the stereotyped group and, in the area of schooling, probably affects confident students more than unconfident ones . . . Stereotype threat should have its greatest effect on the better, more confident students in stereotyped groups, those who have not internalized the group stereotype to the point of doubting their own ability and have thus remained identified with the domain—those who are the academic vanguard of their group.”).

78. “Domain identification” refers to individuals who strongly identify with, and care about performing well in, a particular domain, such as sports or academics. See Nguyen & Ryan, *supra* note 64, at 1315-16.

79. See STEELE, *supra* note 15, at 56-59 (discussing a study examining the effect of stereotype threat on high and low achieving high school students).

80. Steele uses “over-efforting” to describe the desire to disprove the stereotype as an “additional task” capable of producing a “stressful and distracting” form of multitasking. *Id.* at 110-11; Steele & Aronson, *supra* note 34, at 808 (“Black participants . . . reread the test items more than White participants. Such findings do not fit the idea that these participants underperformed because they withdrew effort from the experiment.”); David A. Nussbaum & Claude M. Steele, *Situational Disengagement and Persistence in the face of Adversity*, 43 J. EXPERIMENTAL SOC. PSYCHOL. 127 (2007) (observing that when under threat, Black participants burdened themselves by spending more time on difficult problems than their white counterparts). Accord O’Brien & Crandall, *supra* note 75 (observing that over-efforting exists whenever a negative stereotype is made relevant, but only inhibits performance when the activity pushes participants to the limits of their ability).

Second, self-awareness of one's identity and the latent salience of a relevant negative stereotype in a particular domain are sufficient to trigger stereotype threat.⁸¹ As a result of this "stigma consciousness,"⁸² the negative stereotype associated with an individual's group is readily available and relevant to their performance.⁸³ Within such environments, explicitly priming a stereotype-relevant identity is not always necessary to trigger stereotype threat.⁸⁴ The simple self-awareness that negative stereotypes attach to an individual and her respective performance in a given domain is sufficient. Further, stigma consciousness is most likely to materialize in environments characterized by dramatic underrepresentation. Law school classrooms and testing environments, in which racial underrepresentation and historical stereotypes pervade, are thus ripe for stereotype threat.

C. Challenging the Science—Ecological Validity

Although laboratory studies have conclusively produced stereotype threat effects,⁸⁵ policy decisions ultimately depend on stereotype threat's ecological validity.⁸⁶ Unfortunately, ethical⁸⁷ and pragmatic⁸⁸ obstacles inherent to stereotype threat

81. See *infra* Part II.D.

82. See Elizabeth C. Pinel, *Stigma Consciousness: The Psychological Legacy of Social Stereotypes*, 76 J. PERSONALITY SOC. PSYCHOL. 114 (1999) (employing "stigma consciousness" to refer to the level at which individuals are self-conscious of their stigmatized status).

83. In a retrospective piece, then Trinity College senior Ronald Satterthwaite recalled his own self-awareness: "As I grew older, I became more aware of the negative stereotypes that came with being a black male as far as intelligence and schooling are concerned. I learned that on standardized tests that I was taking, I wasn't supposed to be scoring as high as I did and that, eventually, I had a 20 percent chance of dropping out of school and ending up in jail." Ronald Satterthwaite, *Rising Above the Stereotype Threat, Race, Ethnicity and Me*, 1 (Trinity University Publication).

84. See Steele & Aronson, *supra* note 34, at 808 (concluding that the "mere cognitive availability of the racial stereotype is enough to depress Black participants' intellectual performance").

85. Even skeptics of stereotype threat's ecological validity affirm the existence of stereotype threat within laboratory settings. See, e.g., Paul R. Sackett et al., *On Interpreting Stereotype Threat as Accounting for African American-White Differences on Cognitive Tests*, 59 AM. PSYCHOL. 7 (2004) ("Steele and Aronson clearly showed that imposing and eliminating stereotype threat can, in laboratory settings, affect the test performance of both African American and White students.").

86. Ecological validity refers to the generalizability of a phenomenon beyond the laboratory to real world settings.

87. See Claude M. Steele, Stephen J. Spencer & Joshua Aronson, *Contending with Group Image: The Psychology of Stereotype and Social Identity Threat*, in 34 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY 379 (M. Snyder, ed., 2002) ("Doing something that might affect a person's performance on a real-life high stakes test is, in most imaginable situations, unacceptable without informed consent, which may be unreliable to get in such situations.").

88. It is nearly impossible to measure the level of ambient threat experienced by vulnerable individuals prior to external manipulations. Thus, researchers must make assumptions about whether the baseline reflects the threat or control conditions from laboratory settings. See Lawrence J. Stricker & William C. Ward, *Stereotype Threat, Inquiring About Test Takers' Ethnicity and Gender, and Standardized Test Performance*, 34 J. APPLIED SOC. PSYCHOL. 665, 690 (2004) ("A clear limitation of these studies was that data were only available about test performance and not about its possible causes . . . This limitation is close to inevitable in field experiments in operational settings that must rely on unobtrusive observations."); Michael J. Cullen, Chaitra M. Hardison, & Paul R. Sackett, *Using SAT-Grade and Ability-Job Performance Relationships to Test Predictions Derived from Stereotype Threat Theory*, 89 J. APPLIED PSYCH. 220 (2004) ("In the laboratory setting, it is feasible to present differing cover stories about the purpose of a test . . . In high-stakes testing, the purpose of the test is clear to the applicants and it is doubtful that it would be believable—or ethical—to

constrain our ability to directly measure its relevance beyond the laboratory setting. As a result, few researchers have taken up this task. Among those who have, some have cautioned against generalizing stereotype threat findings to high stakes testing environments.⁸⁹

These critiques take one of two forms: 1) analyzing preexisting data sets in which researchers do not manipulate actual testing environments, or 2) externally manipulating actual testing environments. Michael Cullen, professor of psychology from the University of Minnesota, led a pair of studies representative of the former approach.⁹⁰ To test the generalizability of stereotype threat, Cullen analyzed two massive data sets by comparing the “SAT-grade relationships in the college setting” and “test-job performance relationships in a military setting.”⁹¹ Around the time of Cullen’s first study, Lawrence Stricker and William Ward of the Educational Testing Service undertook a project representative of the latter approach. Stricker and Ward attempted to manipulate the level of threat in testing environments through the strategic inclusion of pre-test questionnaires inquiring about race or gender.⁹²

Regardless of their distinct methods, the researchers came to the similar conclusion that high stakes testing environments failed to produce the same stereotype threat effect observed in laboratory settings.⁹³ To explain this outcome, the researchers extended various hypotheses.⁹⁴ One common hypothesis was that if stereotype is present in laboratory and applied settings, the motivation present in high-stakes environments enables students to overcome otherwise debilitating effects.⁹⁵

While these findings may cause us to pause, several studies reveal the fragility of the conclusions drawn from the foregoing analyses.⁹⁶ The most notable counter-

present information attempting to minimize threat by downplaying the importance of the test or by presenting factually untrue information suggesting no subgroup differences.”).

89. The genesis of these critiques emerged in a series of 2003 articles comprising four independent studies that had attempted to determine the ecological validity of stereotype threat. See Claude M. Steele & Paul G. Davies, *Stereotype Threat and Employment Testing: A Commentary*, 16 HUM. PERFORMANCE 311 (2003). Each article concluded that “more research is needed.” James L. Farr, *Introduction to the Special Issue: Stereotype Threat Effects in Employment Settings*, 16 HUM. PERFORMANCE 179 (2003).

90. See Cullen, Hardison, & Sackett, *supra* note 88; Michael J. Cullen, Shonna D. Waters & Paul R. Sackett, *Testing Stereotype Threat Predictions for Math-Identified and Non-Math-Identified Students by Gender*, 19 HUM. PERFORMANCE 421 (2006).

91. The first data set included SATM score, SATV score and GPA of freshmen from thirteen universities. The second came from the U.S. Army Selection and Classification Project, which included measurements of performance on 10 subtests.

92. See Stricker & Ward, *supra* note 88.

93. See Cullen, Hardison & Sackett, *supra* note 89 (“[S]tereotype threat [was not] an appreciably important explanatory mechanism for the Black-White differences in cognitive ability test scores.”); Stricker & Ward, *supra* note 88 (concluding that the pre-test questionnaire failed to produce statistically significant variation in performance).

94. Cullen hypothesized that salience of stereotype-relevant identity necessary to produce stereotype threat was not present in applied settings. Reflecting their assumption that the pre-test questionnaire sufficiently primed domain-relevant identities, Stricker and Ward conceded that high stakes environments may already contain sufficient threat, such that the impact of the questionnaires was superfluous. Stricker & Ward, *infra* note 102, at 686.

95. *Id.* at 686. *But see* Aronson et al., *supra* note 59.

96. Several studies claim to have observed stereotype threat effects in applied settings. See Jelte Wicherts, Conor V. Dolan & David J. Hessen, *Stereotype Threat and Group Differences in Test Performance: A Question of*

study comes from Professors Nguyen and Ryan, who analyzed over ten years of research on stereotype threat.⁹⁷ In finding that contextual factors moderate the impact of threat,⁹⁸ Nguyen and Ryan concluded that the entirety of the evidence suggests stereotype threat impacts performance in applied settings.⁹⁹ Translating this finding to the SAT, they asserted that stereotype threat would under-measure the talent of a minority student with average cognitive ability by about 50 points.¹⁰⁰

In a separate article, Professors Danaher and Crandall directly responded to Stricker and Ward's analysis and conclusions.¹⁰¹ Danaher and Crandall re-examined the data and concluded that a slight change in methodology revealed a significant stereotype threat effect.¹⁰² They claimed that even with a small stereotype threat effect size, the impact would be great. Contrary to Stricker and Ward, they predicted that removing this additional burden could have led to a 5.9% increase in the number of female students achieving a passing calculus score.¹⁰³

Measurement Invariance, 89 J. PERSONALITY & SOC. PSYCHOL. 696 (2005); Catherine Good, Joshua Aronson, & Jayne Anne Harder, *Problems in the Pipeline: Stereotype Threat and Women's Achievement in High-Level Math Courses*, 29 J. APPLIED DEVELOPMENTAL PSYCHOL. 17 (2007) ("The pattern of results suggests that even among the most highly qualified and persistent women in college mathematics, stereotype threat suppresses test performance."); MITCHELL CHANG ET AL., STEREOTYPE THREAT: UNDERMINING THE PERSISTENCE OF RACIAL MINORITY FRESHMEN IN THE SCIENCES 30-31 (2009) (observing stereotype threat effects among highly domain identified underrepresented minority students who experience negative racial experiences).

The anti-generalizability studies also problematically assume that stereotype-relevant identities are not salient in high stakes settings absent explicit manipulation. See *infra* Part II; Claude M. Steele & Joshua Aronson, *Stereotype Threat Does not Live by Steele and Aronson (1995) Alone*, 59 AM. PSYCH. 47, 48 (2004) ("[T]he stereotype threat conditions, and not the no-threat conditions . . . produce group differences most like those of real-life testing. Stereotype threat conditions represent the test as ability diagnostic, either en passant or by saying nothing at all and relying on participants to know a test when they see one. It is the no-threat conditions that are unlike real-life testing. They present the tests nondiagnostic of the participants' ability or of their group's ability—in stark contrast to real life testing situations."); Walton & Spencer, *supra* note 31, at 1133 ("Because such work does not remove psychological threat, it tests only whether, in control conditions, one measure is completed in more threatening circumstances and is therefore more biased than the other."); Wicherts, Dolan & Hessen, *supra*, at 697 (challenging the assumptions underlying Cullen's conclusions).

97. See Nguyen & Ryan, *supra* note 64.

98. For instance, Nguyen and Ryan found that stereotype threat effects are more severe vis-à-vis racial stereotypes than gender stereotypes. *Id.* at 1327.

99. *Id.*

100. *Id.* at 1330. Standard deviation for the SAT is around 100 points. See http://www.collegeboard.com/prod_downloads/about/news_info/cbsenior/yr2004/2004_CBSNR_total_group.pdf.

101. Compare Kelly Danaher & Christian S. Crandall, *Stereotype Threat in Applied Settings Re-Examined*, 38 J. APPLIED SOC. PSYCHOL. 1639 (2008) (questioning whether the pre-test questionnaire was sufficient to independently make race salient in the high stakes testing environment), with Lawrence J. Stricker & William C. Ward, *Stereotype Threat in Applied Settings Re-Examined: A Reply*; 38 J. APPLIED SOC. PSYCHOL. 1656 (2008) (responding to a critique of the methods used in an earlier study refuting the claim that removing questions about gender or ethnicity on high-stakes tests would remove stereotype threat).

102. See Danaher & Crandall, *supra* note 101, at 1643 ("Stricker and Ward . . . did not accept any effect unless it showed $p < .05$ in the overall ANOVA; $p < .05$ for planned comparisons . . . and also showed $\eta \geq .10$ or $d \leq .20$. By these criteria, the manipulations of holding off inquiries about race or gender until after the test had no effect for the AP Calculus AB test or for any of the CPT scales. However, if one uses the traditional $p < .05$ for the overall ANOVA . . . and employs $\eta \geq .05$ as a standard, we find several important effects of the manipulation in Stricker and Ward's data.").

103. Nguyen argues that when combined with her meta-analysis, these findings suggest that vulnerable

II. THE RECOMMENDATION

The science suggests that vulnerable, stereotype-threatened students experience a measurement bias on the LSAT.¹⁰⁴ Institutions moved by these findings now have the “critical task . . . to determine how to account for this bias so as to make selection decisions that are meritocratic and that do not discriminate against deserving people from stereotyped groups.”¹⁰⁵ In other words, having recognized that our “objective processes” for evaluating merit are infected with a race-dependent mismeasurement, how do we achieve fair measures that embody our color-blind ideal? I offer “rescaling” as the prescription for this ailment. Rescaling is a race-conscious intervention that corrects the mismeasured LSAT¹⁰⁶ scores of stereotype-threatened students. I open this Part with a description of rescaling. Following this introduction, I respond to likely policy objections. Part III supplements this policy analysis with a constitutional defense of rescaling.

A. Rescaling

An accurate understanding of rescaling demands defining the precise contours of the underlying harm: stereotype threat-induced measurement bias. Stereotype threat is neither past discrimination nor present structural or institutional discrimination.¹⁰⁷ Rather, in causing a quantifiable, race-dependent under-measurement of student talent, stereotype threat is identifiable discrimination in the here and now. By relying on test scores that have been skewed by stereotype threat, admissions offices use a metric that undervalues the merit of talented Black and Latino/a students. This race-dependent mismeasure contravenes the goal of a neutral, color-blind, and individualized review of each applicant.¹⁰⁸ Moreover, the harm of this process defect is not limited to those students denied admission as a result of mismeasurement. In systematically excluding qualified Black and Latino/a students, law schools undermine the attainment of diverse student bodies. The few Black and Latino/a students who pass through the gates consequently face additional and unique race-dependent burdens associated with their depleted numbers and token status.¹⁰⁹

test takers are likely to experience an underperformance in high-stakes testing environments. *See* Nguyen & Ryan, *supra* note 64, at 1330.

104. *See* Walton & Spencer, *supra* note 30, at 1137 (“[S]tandard measures of academic performance underestimate the ability and potential of ethnic minority students and of women in quantitative fields.”).

105. *Id.*

106. Law schools may soon have the option to terminate reliance on the LSAT. The American Bar Association (ABA) currently requires the LSAT for law school accreditation. However, a recent report citing problems associated with the LSAT suggests that this requirement may change. *See* ABA May Drop LSAT Requirement, Jan. 14, 2011, available at: http://www.insidehighered.com/news/2011/01/14/aba_may_end_requirement_that_law_schools_use_lsar. Considering the interests motivating law school decisionmaking, universal abandonment of the LSAT is unlikely. *See* West-Faulcon, *supra* note 29 (arguing the law school rankings incentivize reliance on LSAT scores).

107. I am not arguing that institutions should ignore these broader mechanisms of subordination and exclusion. However, as identifiable racial discrimination in the here and now, stereotype threat-produced measurement bias reflects the particular harm recognized by modern antidiscrimination law. *See infra* Part III.

108. *See id.*

109. *See infra* Part II.C.

Additionally, tokenism produces key ingredients necessary to trigger stereotype threat within the law school. Thus, correcting this measurement bias not only remedies a fair measures defect at the moment of admission, but also prevents the vicious cycle that unwarranted exclusion perpetuates.

As a remedy for stereotype threat-induced measurement bias, law schools that adopt my rescaling proposal would add four points to the LSAT score of all Black applicants with a score of 150 and above. Considering that stereotype threat affects Black and Latino/a applicants, it may seem odd that my proposal only engages the mismeasurement of Black students. I articulate rescaling solely in terms of Black applicants in order to maximize the clarity of the proposal. Institutions adopting a rescaling policy would be able to implement a mechanism that corrects the scores of Latino/a applicants in addition to Black applicants.¹¹⁰ In general, rescaling has three defining features. First, it is expressly race-conscious. Since the underlying measurement bias is race-dependent, effectively targeting susceptible students demands a racially attentive policy.¹¹¹ Second, rescaling entails a specific numerical adjustment that corrects the measurement bias associated with stereotype threat. This correction produces a more accurate account of student merit. The 4 points correspond to the mean effect size of stereotype threat on vulnerable Black students.¹¹² Third, rescaling utilizes a 150-point floor. This component of the policy arises from the recognition that stereotype threat-induced underperformance necessitates a baseline ability. The floor thus ensures that rescaling is limited to students suffering from a measurement bias.¹¹³

The following numbers add texture to rescaling. In 2008-2009, Black students had a mean LSAT score of 142 (SD = 8.40).¹¹⁴ In contrast, White students had a mean score of 152 (SD = 8.96).¹¹⁵ Color-blind admissions policies that rely on the LSAT inevitably reproduce this disparity through the production of a student body characterized by Black underrepresentation.¹¹⁶ While stereotype threat may not account for this entire performance gap, the best measure predicts that vulnerable

110. Since stereotype threat does not equally impact individuals from all negatively stereotyped groups, the precise rescaling recommendation will potentially differ depending on the particular racial identity.

111. For a critique of race neutral alternatives, see *infra* Part III.C.4.

112. The four point recommendation derives from personal conversations with Professor Rachel Godsil, Seton Hall School of Law, concerning the results of a study that observed the mean effect size of stereotype threat on MCAT-takers. Extrapolating the mean effect size from the MCAT to the LSAT suggests a minimum mean effect of four points for vulnerable Black LSAT-takers. Since researchers have yet to measure the impact of stereotype threat on LSAT-takers, this is the most analogous measure currently available.

113. I elected a 150-point floor because it corresponds to one standard deviation above the mean LSAT score for Black students. In other words, I am using one standard deviation as a proxy for students in the vanguard of the group. While some may argue that this number is arbitrary, it is likely that the floor is too high and some students scoring below a 150 still suffer from stereotype threat.

114. See *supra* note 21.

115. *Id.*

116. It is arguable that “color-blind” admissions policies have always taken race into account. In some instances, this occurs by granting hard points on a “race-neutral” basis that carry a “race-dependent” functional effect. See Tim Wise, *Whites Swim in Racial Preference*, AlterNet (Feb. 20, 2003), available at <http://www.alternet.org/story/15223/> (illustrating how the University of Michigan’s undergraduate application allotted points along a “race-neutral” axis that in effect could only be attained by white students).

Black LSAT-takers experience a four-point undermeasure of talent.¹¹⁷ Thus, the average Black student scoring a 158 has the same talent as the average White student scoring a 162. We can think of these four points as a direct manifestation of the “pioneer tax” we encountered in Part I.¹¹⁸ In the zero-sum context of law school admissions, this four-point tax on Black students functions as an informal, unspoken four-point boost for White students. If formally contained within an institution’s admissions criteria, such a policy would be universally condemned on policy and constitutional grounds. Considering our hostility toward such a system, why should we find an informal, yet functionally equivalent “tax” to be any more acceptable? Rescaling simply eliminates this tax.

Consider the following hypothetical: Imagine we live in a truly color-blind world, wholly devoid of racial contingencies. Race is nothing more than phenotype, a characteristic that plays no role in an individual’s access to resources or susceptibility to race-dependent psychological barriers. Now consider the constitutional implications of several different admissions policies. World One: Law School X adds four points to the LSAT scores of all White applicants. Is this action constitutional? Clearly not. World Two: Law School X gives a random fifty percent of White applicants four additional points on the LSAT. Constitutional? Again, indisputably not. World Three: Law School X grants a random fifty percent of White applicants a random bump from zero to four points on the LSAT. Constitutional? This still presents us with a formal distribution of points awarded solely on the basis of race, which in a world devoid of racial contingencies constitutes an unconstitutional act. World Four: One final iteration. Law School X grants a random zero to four-point bump to all White applicants with an LSAT score of 150 or higher. Constitutionally, has anything significant changed between World One and World Four? Law School X still awards points on the basis of race. In every world, color-blindness has been compromised.

What is the relevance of these hypotheticals? The four worlds, which frame stereotype threat from the perspective of the White beneficiary, cleanly map onto our best understandings of this phenomenon. World Four resembles contemporary practices.¹¹⁹ The random provision of points and the 150-point baseline reflects the understanding that stereotype threat will not affect all individuals from a negatively stereotyped group equally. In this way, World Four provides a more accurate depiction of contemporary practices than the color-blind ideal many of us espouse. And we already recognized that World Four is not easily distinguishable from World One, where we saw a blatant preference for all White students. If this is the world in which we live, how could it not compel a remedy? And if it does not compel a remedy, does the Constitution prohibit one because it is race-conscious?

The inability to draw clean distinctions between World One and World Four also

117. See *supra* note 112.

118. See *supra* note 74.

119. Considering four points corresponds to the mean effect size of stereotype threat on Black applicants, see *supra* note 112, a more accurate account would have the random allocation of points exceed the four-point ceiling currently in the hypothetical.

calls into question traditional narratives concerning race, merit, and the status quo.¹²⁰ Kang and Banaji distill this concern, suggesting that it “should give us pause as we confront the fact that arbitrary environmental cues can trigger implicit cognitive processes that interfere with or facilitate performance on seemingly objective measures.”¹²¹ What “we thought to be fair assessments of ‘merit’ . . . turn out to be [racially tinged] mismeasurements—not because of explicit animus but because of hidden mental processes that by their nature cannot reach conscious awareness.”¹²² This means that the LSAT, even if administered in a uniform fashion to all students, fails as a neutral measure of talent.¹²³

The harm flowing from race-dependent mismeasures is not limited to the moment of decision making. Rather, reliance on biased measures topples the first domino in a vicious cycle. Notwithstanding stereotype threat, a small number of Black and Latino/a students are still able to gain admission. However, as a result of the disproportionate exclusion of qualified Black and Latino/a students, those who enter face additional burdens unique to token status.¹²⁴ Beyond independent challenges associated with severe underrepresentation, tokenism has the potential to trigger the same stereotype threat effects that produced the initial mismeasures. Thus, process defects at the moment of decisionmaking sew the ground for subsequent race-dependent obstacles that majoritarian students never encounter.

B. Policy Objection: Overprediction

One inevitable objection to rescaling is the claim of overprediction.¹²⁵ Overprediction is the notion that standardized tests actually overpredict the performance of

120. See generally Susan Sturm & Lani Guinier, *The Future of Affirmative Action: Reclaiming the Innovative Ideal*, 84 CALIF. L. REV. 953 (1996) (“Competing narratives drive the affirmative action debate . . . Each story proceeds from different assumptions about the baseline of decision making: how fair, unbiased, and merit-driven is the system in which affirmative action operates?”).

121. Kang & Banaji, *supra* note 13, at 1089. See also William C. Kidder, *Portia Denied: Unmasking Gender Bias on the LSAT and Its Relationship to Racial Diversity in Legal Education*, 12 YALE J.L. & FEMINISM 1, 26 (2000) (“Stereotype threat research suggests that the LSAC’s simplistic definition of standardized conditions obscures how a history of sexism, racism and classism can facilitate so-called standardized conditions which further privilege White and male and affluent (and therefore especially affluent White male) test takers.”).

122. Kang & Banaji, *supra* note 13, at 1089.

123. See STEELE, *supra* note 15, at 60 (“The reality of stereotype threat also made the point that places like classrooms, university campuses, standardized-testing rooms, or competitive-running tracks, though seemingly the same for everybody, are, in fact, different places for different people.”); See also Expert Report of David M. White, *Grutter v. Bollinger*, 16 F. Supp. 2d 797 (E.D. Mich. 1998) (No. 97-75928) at 17 (“Nevertheless, the different ways students take tests can affect their scores without reflecting their underlying ability to read or reason. Simply giving the same test to all applicants cannot ensure that all applicants will take the same test.”).

124. See *infra*, Part II.C.

125. Another likely critique would follow the logic of mismatch theory. Mismatch theory posits that race-conscious admissions policies counter-intuitively reduce the number of Black law students who eventually become practicing attorneys. See Richard H. Sander, *A Systematic Analysis of Affirmative Action in American Law Schools*, 57 STAN. L. REV. 367, 474-77 (2004) (“The number of black lawyers produced by American law schools each year and subsequently passing the bar would probably increase if those schools collectively stopped using racial preferences . . . the absolute number of black law graduates passing the bar on their first attempt . . . would be much larger under a race-blind regime than under the current system of preferences.”).

negatively stereotyped groups.¹²⁶ Contrary to the claim underlying this Article, overprediction asserts that the LSAT actually inflates minority talent.¹²⁷ This claim stems from the observation that when controlling for entering indicia of ability, as measured by the LSAT, well prepared and highly motivated Black students fare worse than their White counterparts in future academic performance.¹²⁸ While the ubiquity of this phenomenon is evident, it is important to note that recent reports have begun to question whether overprediction accurately describes minority performance.¹²⁹

Even accepting the claim of overprediction, there are multiple ways to interpret

For a detailed critique of mismatch theory, see andre douglas pond cummings, “Open Water”: *Affirmative Action, Mismatch Theory and Swarming Predators: A Response to Richard Sander*, 44 BRANDEIS L.J. 795, 799 n.22, Part III (2006). See also Cheryl I. Harris & William C. Kidder, *The Black Student Mismatch Myth in Legal Education: The Systematic Flaws in Richard Sander’s Affirmative Action Study*, J. BLACKS IN HIGHER EDUC. 102, 103 (2004/5), (relying on more typical data from 2003 and 2004 reveals that eliminating “affirmative action would slash African American law school admissions by 24 percent and 33 percent, respectively”); Jesse Rothstein & Albert Yoon, *Affirmative Action in Law School Admissions: What do Racial Preferences Do?*, 75 U. CHI. L. REV. 649, 707 (2008) (“The number of black law graduates would fall by 60 percent, while the numbers of bar entrants and large-firm associates would each fall by half.”); David L. Chambers et al., *The Real Impact of Eliminating Affirmative Action in American Law Schools: An Empirical Critique of Richard Sander’s Study*, 57 STAN. L. REV. 1855 (2005); Ian Ayres & Richard Brooks, *Does Affirmative Action Reduce the Number of Black Lawyers?*, 57 STAN. L. REV. 1807, 1809 (2005) (utilizing the same data sets as Sander to conclude that eliminating race-conscious admissions programs would result in a net reduction of Black attorneys); Linda F. Wightman, *The Threat to Diversity in Legal Education: An Empirical Analysis of the Consequences of Abandoning Race as a Factor in Law School Admission Decisions*, 72 N.Y.U. L. REV. 38 (1997). Sander’s argument is a contemporary reincarnation of past “mismatch” theses promulgated in opposition to affirmative action. See Harris & Kidder, *supra*.

126. See, e.g., STEVEN FARRON, *THE AFFIRMATIVE ACTION HOAX* (2005); Gail L. Heriot & Christopher T. Wonnell, *Standardized Tests Under the Magnifying Glass: A Defense of the LSAT Against Recent Charges of Bias*, 7 TEX. REV. L. & POL. 467 (2002/3); Leonard Ramist et al., *Using Achievement Tests/SAT II: Subject Tests to Demonstrate Achievement and Predict College Grades: Sex, Language, Ethnic and Parental Education Groups*, College Board Research Report No. 2001-5 (2001).

127. Overprediction assumes that Black and Latino/a students are less qualified than Whites with identical LSAT scores. See Rogers Elliott et al., *The Role of Ethnicity in Choosing and Leaving Science in Highly Selective Institutions*, 37 RES. HIGHER EDUC. 681, 684 (1996) (“At White-majority institutions non-Asian minorities are, by virtue of race-preferential admission policies, at an often serious disadvantage with respect to validly predictive indices of talent, and if equally developed ability predicts equal persistence, unequally developed ability should predict differential persistence.”).

128. See e.g., JAMES EARL DAVIS, *CAMPUS CLIMATE, GENDER, AND ACHIEVEMENT OF AFRICAN AMERICAN COLLEGE STUDENTS* 3 (1998) (“[T]oo many [Black students] exhibit a marked decrease in performance from their high school grades over and beyond what is generally expected for adjustment to college level work.”); Steele & Aronson, *supra* note 33, at 798 (“At every level of preparation as measured by a standardized test—for example, the Scholastic Aptitude Test (SAT)—Black students with that score have poorer subsequent achievement—GPA, retention rates, time to graduation, and so on—than White students with that score.”).

129. See, e.g., LAW SCHOOL ADMISSION COUNCIL, *LSAT TECHNICAL REPORT SERIES: ANALYSIS OF DIFFERENT PREDICTION OF LAW SCHOOL PERFORMANCE BY RACIAL/ETHNIC SUBGROUPS BASED ON 2005-2007 ENTERING LAW SCHOOL CLASSES* (2009), available at <http://www.lsac.org/LSACResources/Research/TR/TR-09-02.pdf>. The results “provided no evidence that LSAT scores . . . unfairly predict future law school performance for any racial/ethnic group.” *Id.* See also Rothstein & Yoon, *supra* note 125, at 711-12 (“There is little evidence of underperformance among black students with entering LSAT scores and undergraduate GPAs above those of the twentieth-percentile . . . [T]here is no evidence of black underperformance on any employment outcomes.”).

these findings.¹³⁰ One is to assume that standardized tests overpredict Black and Latino/a talent. An alternative explanation is that once in college or law school, something prohibits Black and Latino/a students from reaching their full potential.¹³¹

C. Alternative Explanations: Tokenism and Racially Hostile Environments

“The problem is not so much the entry; it’s what happens while you’re there . . .”¹³²

As with other institutions, the University of Michigan Law School observed that its minority students underperformed relative to their White counterparts.¹³³ Michigan had recognized that underperformance was partly due to the lack of a “critical mass”¹³⁴ of minority students. Absent a critical mass of non-White students, educational spaces such as law school classes have the tendency to produce racially hostile environments,¹³⁵ which impose race-dependent burdens on minority students.¹³⁶ Similar to stereotype threat,¹³⁷ these race-dependent burdens uniquely hinder the

130. See, e.g., Walton & Spencer, *supra* note 30, at 1133 (arguing that overprediction may exist not from overmeasurements, but rather because the “the level of psychological threat increases at each rung of the educational ladder, for instance as students become more anonymous and as they reach the edge of their abilities”). See also, Gregory M. Walton & Geoffrey L. Cohen, *A Question of Belonging: Race, Social Fit, and Achievement*, 92 J. PERSONALITY & SOC. PSYCHOL. 82 (2007).

131. If using SAT or LSAT scores as a reference, it is likely these scores already involve measurement defects. Thus, a more accurate measure of “potential” would be a rescaled score, in which stereotype threat is taken into account.

132. Katherine S. Mangan, *Does Affirmative Action Hurt Black Law Students?*, Chron. Higher Educ., Nov. 12, 2004, at 35 (quoting Harvard Law School graduate Kathy Hart, who is Black).

133. See *Grutter v. Bollinger*, 539 U.S. 306 (2003).

134. Critical mass refers to the “concentration of a ‘meaningful’ number of underrepresented students necessary to create an environment in which such students can fully engage in the classroom as individuals rather than felling like they have to be a spokesperson for their race or defy stereotypes.” Deidre M. Bowen, *Brilliant Disguise: An Empirical Analysis of a Social Experiment Banning Affirmative Action*, 85 IND. L.J. 1197, 1199. n.7 (2010).

135. Inzlicht & Ben-Zeev refer to “threatening intellectual environments” to describe spaces that “can activate the threatening effects of gender stereotypes.” Michael Inzlicht & Talia Ben-Zeev, *A Threatening Intellectual Environment: Why Females are Susceptible to Experiencing Problem Solving Deficits in the Presence of Males*, 11 PSYCHOL. SCI. 365, 365 (2000). My use of the term connotes the same meaning, yet with respect to race. Allen and Solorzano have identified fourteen characteristics of negative racial climates that exist in the absence of a critical mass. Examples include: “Students of color experience numerous overt instances of race discrimination; however, the negative campus racial climate . . . actually originates from more subtle, covert racial incidents”; and “The negative campus racial climate exacts psychological and behavioral tolls on students of color that interfere with their academic achievement.” Walter R. Allen & Daniel Solorzano, *Affirmative Action, Educational Equity and Campus Racial Climate: A Case Study of the University of Michigan Law School*, 12 BERKELEY LA RAZA L.J. 237, 299 (2001). For a more detailed description of critical mass and racially hostile environments, see Elizabeth Chambliss & Christopher Uggen, *Men and Women of Elite Law Firms: Reevaluating Kanter’s Legacy*, 25 LAW & SOC. INQUIRY 41, 43-48 (2000); Sylvia Hurtado et al., *Enacting Diverse Learning Environments: Improving the Climate for Racial/Ethnic Diversity in Higher Education*, ASHE-ERIC Higher Education Report 26 (1999); Emily Calhoun, *An Essay on the Professional Responsibility of Affirmative Action in Higher Education*, 12 TEMP. POL. & CIV. RTS. L. REV. 1, 2-5 (2002).

136. See Allen & Solorzano, *supra* note 135, at 246 (observing that undergraduate students of color felt that their colleges are racially charged atmospheres wherein their presence is questioned, yet the presence of Whites on campus is not questioned).

137. See Claude Steele & Joshua Aronson, *Stereotype Threat and the Intellectual Test Performance of African Americans*, 69 J. PERSONALITY & SOC. PSYCHOL. 797, 810 (1995) (“Although probably in the same family of

ability of qualified and talented token students to reach their potential.¹³⁸ Furthermore, majority students never have to bear these burdens.¹³⁹

Professor Timothy Clydesdale describes the divergent experiences of token and majority students as “a forked river, with one side offering a challenging and often dangerous white-water rapids course, the other side a swift but smooth current, and an entry gate that variously restricts access to the river.”¹⁴⁰ Three characteristics unique to tokens facilitate this divergence:¹⁴¹ super-visibility,¹⁴² polarization,¹⁴³ and distortion of individual characteristics.¹⁴⁴ As relative underrepresentation rises, the debilitating effects of these characteristics increase.¹⁴⁵ While these burdens occur across a variety of settings, spaces that maintain pervasive negative stereotypes about tokens are the most likely to produce racially hostile environments. Well-known examples include the following: Black males in traditionally White colleges;¹⁴⁶ under-

effects as stereotype threat, token status would be expected to disrupt cognitive functioning even when the token individual is not targeted by a performance-relevant stereotype, as with, for example, a White man in a group of women solving math problems. Nor do stereotype threat effects require token status . . . [i]n real life, of course, these two processes may often co-occur, as for the Black in an otherwise non-Black classroom. They are nonetheless, distinct processes.”)

138. See, e.g., Charles G. Lord & Delia S. Saenz, *Memory Deficits and Memory Surfeits: Differential Cognitive Consequences of Tokenism for Tokens and Observers*, 49 J. PERSONALITY & SOC. PSYCHOL. 918 (1985) (observing the token status can produce short term memory strain).

139. See Brian Owsley, *Black Ivy: An African-American Perspective on Law School*, 28 COLUM. HUM. RTS. REV. 501, 524 (1997) (“It’s funny how whites can turn on and off this whole [racial] dialogue thing, but once I get started, I’m screwed up and freaked out by it I can’t concentrate or study.”)

140. Timothy Clydesdale, *A Forked River Run Through Law School: Toward Understanding Race, Gender, Age, and Related Gaps in Law School Performance and Bar Passage*, 29 LAW & SOC. INQUIRY 711, 712 (2004).

141. Contemporary understandings stem from the 1977 work of Rosabeth Kanter, who analyzed the experiences of token women in male-dominated workplaces. See Rosabeth M. Kanter, *Some Effects of Proportions on Group Life: Skewed Sex Ratios and Responses to Token Women*, 82 AM. J. SOC. 965, 966 (1977) (defining tokens as members of groups making up 15% or less of the workforce). Kanter’s hypothesis, which has been extended beyond gender, is that “women in the setting with fewer female colleagues (1) will do less well academically (performance pressure), (2) will be less integrated into the law school (social isolation), and (3) will be more likely to demonstrate traditional female preferences in study strategies and career choices (role entrapment).” Spangler et al., *Token Women: An Empirical Test of Kanter’s Hypothesis*, 84 AM. J. SOC. 160, 162 (1978).

142. See Spangler et al., *supra* note 141, 161 (1978) (explaining that tokens are “over-observed,” creating additional performance pressures unique to minority members).

143. See *id.* (explaining that characteristics distinguishing minority and majority group members gain increased salience, even when irrelevant to task performance, leading to isolation of the minority member).

144. See *id.* (explaining that majority members impose stereotypes on individual tokens, failing to accept behavior that challenges stereotypes).

145. See Inzlicht & Ben-Zeev, *supra* note 135, at 369 (“[P]erformance deficits tend to increase as the relative number of [majority members] increases.”); Spangler et al., *supra* note 142 (finding that performance pressure hindered the performance of women more severely when women represented 20% of the student body than when they represented 33%).

146. See, e.g., James Earl Davis, *College in Black and White: Campus Environment and Academic Achievement of African American Males*, 63 J. NEGRO EDUC. 620 (1994) (observing that exposure to prejudice at predominantly White institutions is particularly pernicious for Black males’ social and academic development); DAVIS, *supra* note 129, at 4 (“Black students at predominantly white colleges report that racial discrimination occurs with much greater frequency there than at other types of institutions.”); JACQUELINE FLEMING, *BLACKS IN COLLEGE: A COMPARATIVE STUDY OF STUDENTS’ SUCCESS IN BLACK AND WHITE INSTITUTIONS* (1984) (observing that exposure to prejudice on campus has a significant effect on black students’ cognitive and affective development at predominantly white institutions).

represented women in law school;¹⁴⁷ minority students in law school;¹⁴⁸ and tokens in elite law firms.¹⁴⁹ It is thus not surprising that law schools, which persist as spaces where Black and Latino/a students continue to be racialized as unqualified and untalented,¹⁵⁰ remain particularly dangerous for underrepresented students of color.

Professor Deidre Bowen recently provided new insight into the relationship between tokenism, hostile environments and higher education.¹⁵¹ She conducted a comparative study analyzing the experiences of Black and Latino/a students in states that allow affirmative action and states that prohibit the practice.¹⁵² While overt acts of racism continue to be widespread, Professor Bowen discovered that Black and Latino/a students are twice as likely to receive discriminatory treatment in states that have banned affirmative action.¹⁵³ Token status was a central factor contributing to the discrimination. In contrast to minority students who had never been solo in a classroom, solo minorities were four times as likely to experience overt racism from other students, and twice as likely to experience such racism from faculty.¹⁵⁴ Mirroring earlier work, Bowen's findings lend empirical support for the claim that as their overall representation falls, tokens face an increase in race-dependent burdens.¹⁵⁵

Stereotype threat and tokenism are integrally connected phenomena. White law students never confront either of these race-dependent burdens. In contrast, stereotype threat and tokenism function as cyclical and mutually reinforcing burdens that trap Black and Latino/a students in a vicious cycle. Beginning with the LSAT,

147. See, e.g., Lani Guinier et al., *Becoming Gentlemen: Women's Experiences at One Ivy League Law School*, 143 U. PA. L. REV. 1 (1994) ("[T]he experience of women in the aggregate differs markedly from that of their males peers.").

148. See, e.g., Brian Owsley, *Black Ivy: An African-American Perspective on Law School*, 28 COLUM. HUM. RTS. REV. 501, 524 (1997); Cecil J. Hunt, *Guests in Another's House: An Analysis of Racially Disparate Bar Performance*, 23 FLA. ST. U. L. REV. 721 (1996).

149. See, e.g., Robin J. Ely, *The Power in Demography: Women's Social Construction of Gender Identity at Work*, 38 ACAD. MGMT. J. 589 (1995). But see KAY DEAUX & JOSEPH ULLMAN, *WOMEN OF STEEL* (1983).

150. See, e.g., Kendra Fox-Davis, *A Badge of Inferiority: One Law Student's Story of a Racially Hostile Educational Environment*, 23 NAT'L BLACK L.J. 98 (2010); Elizabeth M. Iglesias et al., *Labor and Employment in the Academy—A Critical Look at the Ivory Tower*, 6 EMPLOYEE RTS. & EMP. POL'Y J. 129 (2002); Margaret E. Montoya, *Mascaras, Trenzas, y Grenas: Un/Masking the Self While Un/Braiding Latina Stories and Legal Discourse*, 17 HARV. WOMEN'S L.J. 185 (1994). This literature has also discussed the adverse impact of tokenism on professors of color. See, e.g., PATRICIA WILLIAMS, *THE ALCHEMY OF RACE AND RIGHTS* (1991); Paulette Caldwell, *A Hair Piece: Perspectives on the Intersection of Race and Gender*, 1991 DUKE L.J. 365 (1991).

151. See Deidre M. Bowen, *Brilliant Disguise: An Empirical Analysis of a Social Experiment Banning Affirmative Action*, 85 IND. L.J. 1197 (2010). See also Allen Solorzano, *supra* note 135 (completing an empirical study of racial climate at the University of Michigan Law School and its "feeder" institutions).

152. Bowen's work also challenges the common assertion that affirmative action stamps its beneficiaries with a badge of inferiority. Unlike this narrative suggests, racial hostility is most rampant in states that have positively banned race-conscious admissions. See Bowen, *supra* note 151.

153. See *id.*, at 1221-22 tbl.2, 1234 (finding that "[s]tudents who attend schools in anti-affirmative action states find themselves engaged in an unfriendly environment. Despite being admitted on purely white, normative admissions standards, these students were more likely than any other group to encounter . . . open hostility, internal stigma, and external stigma.").

154. *Id.* at 1228-29.

155. See Allen & Solorzano, *supra* note 135, at 301 ("Women and students of color experience these campuses as hostile environments, places where they are either not welcome or are welcome only in clearly delimited, subordinate status.").

stereotype threat produces a measurement bias that undermeasures the talent of Black and Latino/a students. As a result of the disproportionate exclusion of qualified Black and Latino/a students that follows, those who enter must now pay a new, race-dependent “token-tax.”¹⁵⁶ In addition to creating independent challenges, token status increases the salience of race within a law school environment that continues to attach negative stereotypes to students of color. This combination produces high levels of stigma consciousness, and thereby reproduces the conditions ripe for stereotype threat.¹⁵⁷ We have thus completed the cycle. Rescaling, as a means to promote fair measures and effectively mitigate severe underrepresentation, looks more essential than ever.

III. CONSTITUTIONAL ANALYSIS

A. *Rescaling: A Constitutional Challenge*

Considering the traction of color-blindness,¹⁵⁸ an institution that adopts rescaling will face immediate legal challenge. In anticipation of such challenges, the remainder of this Article offers a defense of rescaling’s constitutionality. To limit abstraction, assume that UCLA School of Law implements rescaling.¹⁵⁹ A constitutional challenge will likely mirror those brought against the University of Michigan in 2003.¹⁶⁰ As a state institution, UCLA must demonstrate that rescaling conforms to the Fourteenth Amendment’s Equal Protection Clause.¹⁶¹

The first battle will be fought over the appropriate level of review. Considering the Supreme Court’s hostility toward “racial classifications,”¹⁶² rescaling appears destined for strict scrutiny. Still, I argue that this is a premature concession, as the Court’s equal protection jurisprudence is far less clear than it initially appears. First, the Court has never held that the mere mention of race within a statute or policy categorically triggers strict scrutiny. For example, civil rights laws expressly mention race but have never been subjected to this level of review.¹⁶³ Thus, rescaling’s attentiveness to race is not the end of the discussion, but only the beginning.

156. I intend “token-tax” to parallel Claude Steele’s “pioneer tax.” See Steele, *supra* note 75.

157. See Inzlicht & Ben-Zeev, *supra* note 136, at 365, 369 (“Simply placing high-achieving women in an environment in which men outnumbered them can cause them to experience performance deficits in a stereotype problem-solving domain.”).

158. See *supra* note 10. Additionally, see Ward Connerly’s crusade against affirmative action, available at http://www.acri.org/ward_bio.html.

159. I choose UCLA School of Law because it is my home institution. This is just a hypothetical.

160. See *Grutter v. Bollinger*, 539 U.S. 306 (2003); *Gratz v. Bollinger*, 539 U.S. 244 (2003). Because UCLA resides in California, rescaling opponents would also claim that the policy violates the California Constitution. See CAL. CONST. art.1 § 31. I choose not to address the merits of this challenge for two reasons. First, only four states currently have anti-affirmative action laws. Thus, the argument’s relevance to a national audience is minimal. Second, my defense of rescaling is premised on the notion that the policy actually promotes color-blindness. Thus, rescaling would arguably place UCLA in greater compliance with California state law than it currently sits.

161. U.S. CONST. amend. XIV § 1.

162. See *Grutter*, 539 U.S. at 326 (“We have held that all racial classifications imposed by government ‘must be analyzed by a reviewing court under strict scrutiny.’”).

163. See, e.g., 42 U.S.C. § 1981 (“All persons within the jurisdiction of the United States shall have the

Notwithstanding what the Court often claims, not all racial classifications are subject to strict scrutiny.¹⁶⁴ This inconsistent treatment arises from the normative judgment that racial classifications produce a legally cognizable harm only within certain contexts.¹⁶⁵ It is thus important to distill the rationale underlying the Court's decision to apply strict scrutiny to racial classifications in the domain relevant to rescaling: higher education.¹⁶⁶ In the realms of employment and higher education, the Court's skepticism towards racial classifications arises from the notion that such action represents an inherently preferential repudiation of race-neutral, meritocratic criteria.¹⁶⁷ When viewing race as an arbitrary characteristic,¹⁶⁸ allocating resources on a race-conscious basis appears to contravene color-blindness by favoring "one race over another."¹⁶⁹ From this baseline, the harm arising from the use of racial classifications is the denial of a neutral process and unfair redistribution of scarce resources. The harm is not simply making distinctions on the basis of race.

Well-known cases from the realms of employment and education exemplify this conceptualization of harm. In *Richmond v. Croson*, Justice O'Connor struck down Richmond's thirty percent minority set-aside not because race was present, but because she believed that the program allocated a race-dependent benefit.¹⁷⁰ This sentiment is not new. First emerging in *The Civil Rights Cases*,¹⁷¹ and then reappearing in Justice Powell's *Bakke* opinion, the Court has historically associated racial classifications with preference whenever a public benefit or resource is at stake.¹⁷²

same right . . . and full and equal benefit of all laws and proceedings . . . as is enjoyed by *white* citizens.") (emphasis added); 42 U.S.C. § 1982 ("All citizens . . . shall have the same right . . . as is enjoyed by *white* citizens.") (emphasis added).

164. Racial classifications in the domain of criminal justice are not subject to strict scrutiny. See *Brown v. City of Oneonta*, 235 F.3d 769 (2d Cir. 2000) (holding that police use of racial classifications is not subject to strict scrutiny). See also, Richard Banks, *Race-Based Suspect Selection and Colorblind Equal Protection Doctrine and Discourse*, 48 UCLA L. REV. 1075 (2001).

165. Compare *City of Oneonta*, 235 F.3d at 772 (holding that strict scrutiny in the criminal justice context is inapplicable because it would chill police protection and "likely undermine the strict scrutiny standard itself"), with *Gratz*, 539 U.S. at 270 (citing *Fullilove v. Klutznick*, 448 U.S. 448, 537 (1980)) ("[R]acial classifications are simply too pernicious to permit any but the most exact[ing review].").

166. In the public school integration context, for instance, the mere use of racial classifications is a harm in and of itself. See *Parents Involved v. Seattle Sch. Dist. No. 1*, 551 U.S. 701 (2007).

167. A majority of the Court has never held that the racial classifications produce a legally cognizable harm in these contexts merely from the mention of race. Still, a plurality of the Court firmly espouses this profoundly color-blind view. See, e.g., *Grutter*, 539 U.S. at 353 (Thomas, J., concurring in part and dissenting in part) ("The Constitution abhors classifications based on race, not only because those classifications can harm favored races or are based on illegitimate motives, but also because every time the government places citizens on racial registers and makes race relevant to the provision of burdens or benefits, it demeans us all.").

168. See Neil Gotanda, *A Critique of "Our Constitution is Color-Blind,"* 44 STAN. L. REV. 1 (1991) (articulating the notion of formal race).

169. *Richmond v. Croson Co.*, 488 U.S. 469, 497 (1989).

170. *Id.* ("[R]eview under the Equal Protection Clause is not dependent on the race of those burdened or benefited by a particular classification.").

171. See *The Civil Rights Cases*, 109 U.S. 3, 25 (1883) ("When a man has emerged from slavery . . . there must be some stage in progress of his elevation when he takes the rank of a mere citizen, and ceases to be the special favorite of the laws.").

172. See *Regents of Univ. of Cal. v. Bakke*, 438 U.S. 265 (1978). See also *Grutter*, 539 U.S. at 306; *Gratz v. Bollinger*, 539 U.S. 244 (2003).

Similar to that of Justice O'Connor, Justice Powell's hostility toward racial quotas was not founded upon the mere presence of race. Rather, it arose from the understanding that admissions decisions based on race contravened merit by extending an unjustifiable preference to otherwise unqualified minority applicants.¹⁷³ As a result of this preference, Justice Powell feared that "innocent [White applicants would have had] to bear the burdens of redressing grievances not of their making."¹⁷⁴

This characterization of the policy was not inevitable. However, it arose from the way in which the UC Davis Medical School framed its own program. The medical school justified its policy as a response to societal discrimination, not as a mechanism to achieve more meritocratic admissions standards.¹⁷⁵ If the medical school had taken the latter route, preference concerns may have dissipated.¹⁷⁶ Justice Powell emphasized this point by distinguishing *Bakke* from a series of opinions in which the Court had not applied strict scrutiny to racial classifications.¹⁷⁷ These cases involved policies designed to cure specific identifiable discrimination.¹⁷⁸ The precedent cases thus entailed a close nexus between the remedy and the underlying harm.¹⁷⁹ Due to the absence of such a nexus in *Bakke*, Justice Powell applied strict scrutiny to a policy he viewed as inherently preferential.¹⁸⁰

Following the foregoing reasoning, rescaling need not trigger strict scrutiny. The goal of remedying race-dependent measurement biases implicates race in a way societal discrimination justifications never could. By focusing on societal discrimination, the medical school implicitly accepted that the minority students who would have been excluded but for the challenged policy were underqualified and undeserving of admission (or at least less-deserving than higher scoring White applicants). Thus, the policy never escaped the Court's historic association of race-consciousness with preference. However, rescaling is fundamentally different. It is not founded upon the notion that a lack of resources, motivation, or culture place Black and Latino/a students behind their White counterparts. Rather, rescaling is premised on the notion that stereotype threat produces an undermeasurement of Black and Latino/a talent on admissions exams. This is a repudiation of the tools, not the students. By mitigating a quantifiable, race-dependent process defect at the particular site of mismeasurement, rescaling contains the tight nexus lauded by Justice Powell. Understanding rescaling in this way suggests that such a policy warrants the

173. *Bakke*, 438 U.S. at 298-90 ("[T]here are serious problems of justice connected with the idea of preference itself.").

174. *Id.* at 298.

175. *Id.*

176. *Id.* at 306 n.43.

177. *Id.* at 300-01 (discussing school desegregation, employment discrimination, and sex discrimination cases).

178. *Id.*

179. *Id.* at 301 ("[W]e approved a retroactive award of seniority to a class of Negro truckdrivers who had been the victims of discrimination—not just by society at large, but by the respondent in that case.").

180. *Id.* Justice Powell never discarded his use of "preference" when describing the precedent cases. Even when characterizing the remedies as dealing directly with identifiable discrimination, the notion that a counter-measure was not inherently preferential remained beyond the scope of Justice Powell's opinion.

limited level of judicial skepticism extended in the precedent cases, not the strict scrutiny applied in *Bakke*.

Still, in light of the modern Court's growing hostility toward any racially attentive policy and apparent inability to differentiate between race-conscious programs that dismantle structures of inequality from policies that produce inequality,¹⁸¹ it is unlikely the Court will apply anything less than strict scrutiny to rescaling.¹⁸² In the remainder of this Article, I argue that rescaling withstands even the Court's most rigid review.

B. *Fatal in Fact*:¹⁸³ *Compelling Interest*

To withstand strict scrutiny, the government must demonstrate that the challenged policy is narrowly tailored to achieve a compelling state interest.¹⁸⁴ The Court has identified student body diversity and remedying past discrimination as sufficiently compelling interests.¹⁸⁵ Thus, institutions of higher education no longer face the challenge of establishing this initial prong.¹⁸⁶ The challenge arises in

181. See *Ricci v. Destefano*, 129 S. Ct. 2658 (2009) (suggesting that all government employment decisions that take race into account will be subject to strict scrutiny). For a relevant critique of this form of jurisprudence, see Cheryl I. Harris & Kimberly West-Faulcon, *Reading Ricci: Whitening Discrimination, Racial Test Fairness*, 58 UCLA L. REV. 73 (2010).

182. See *Bakke*, 438 U.S. at 290 ("Racial and ethnic classifications, however, are subject to stringent examination without regard to these additional characteristics."). In her rationale for the application of strict scrutiny with benign racial classification, Justice O'Connor stated that "the purpose of strict scrutiny is to 'smoke out' illegitimate uses of race by assuring that the legislative body is pursuing a goal important enough to warrant use of a highly suspect tool." *Richmond v. Croson*, 488 U.S. 469, 493 (1989).

183. Professor Adam Winkler conducted an empirical analysis to determine the veracity of this now common idiom. See Adam Winkler, *Fatal in Theory and Strict in Fact: An Empirical Analysis of Strict Scrutiny in the Federal Courts*, 59 VAND. L. REV. 793 (2006). His findings demonstrate that "fatal in fact" may be more myth than reality. However, while federal courts have upheld a substantial number of race-conscious employment and admissions policies, these policies have almost universally arisen in response to specific instances of identifiable discrimination or the goal of diversity in higher education. Thus, "fatal-in-fact" remains an accurate description for race-conscious policies that do not fit neatly into these boxes.

184. See *Bakke*, 438 U.S. at 356.

185. See *Parents Involved v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 720-22 (2007).

186. Two additional conceptualizations of the compelling interest underlying rescaling deserve brief mention. The first is an interest in avoiding Title VI liability. See West-Faulcon, *supra* note 30, at 1082 ("[A]ffirmative action-less universities are admitting so few minorities that the racial disparities in admissions to those institutions establishes a rebuttable legal presumption of a Title VI disparate impact claim."). Professor Kimberly West-Faulcon explains that "[a]ssuming the Fourteenth Amendment permits institutions to use race-conscious admissions policies for the remedial purpose of avoiding Title VI liability, in such a circumstance, remedial affirmative action is simply taking race into account [as] equalizing treatment." Cf. Calhoun, *supra* note 136. The second conceptualization comes from Zatz's nonaccommodation theory. See Noah Zatz, *Managing the Macaw*, 109 COLUM. L. REV. 1357, 1389 (2009) ("[N]onaccommodation theory does not require discriminatory intent, that is, internal causation. Instead, an employer may 'discriminate' even though its decisionmaking process is 'neutral' in the sense that it ignores an employee's protected trait."). Nonaccommodation suggests that an institution discriminates against an applicant, even if the admissions process is "neutral," in the sense that it ignores an applicant's protected trait. In this case, ignoring the fact that reliance on the LSAT bestows an unequal burden on vulnerable Black and Latino/a students would qualify as discrimination. Cf. Michael J. Yelnosky, *The Prevention Justification Theory*, 64 OHIO ST. L.J. 1385 (2003) (suggesting the compelling interests underlying race-conscious employment practices designed to prevent workplace discrimination arise from dramatic gender or racial imbalances).

demonstrating that the chosen policy is narrowly tailored to further one of these ends.¹⁸⁷

Traditionally, the Court has understood attaining diversity and remedying discrimination as compelling interests that exist independently of each other. Neither interest alone wholly encapsulates rescaling, which is motivated by the interdependent relationship between fair measures defects and the underrepresentation of minority students. An accurate portrayal of rescaling thus necessitates situating the policy within the dual goals of attaining diversity and remedying discrimination. Without this comprehensive understanding of the state's interests, rescaling becomes vulnerable to narrow tailoring challenges that fail to recognize that the policy is pursuing both goals simultaneously.¹⁸⁸ In the analysis that follows, I demonstrate that the two established compelling interests form a natural union that embodies the dual motivations underlying rescaling. Beyond providing a secure foundation for the narrow tailoring analysis, the marriage of attaining diversity and remedying discrimination produces a hybrid compelling interest that is more persuasive than the sum of its component parts.

Justifying rescaling solely on diversity grounds proves incomplete. While traditional diversity arguments encompass rescaling's goal of achieving a critical mass of underrepresented students,¹⁸⁹ this is the extent of the commonality. Writing for the Court in *Grutter*, Justice O'Connor's articulation of diversity as a compelling interest reflected a severely limited understanding of diversity's core value. According to the Court, diversity provides "educational benefits" by infusing a wide range of ideas and perspectives into the classroom.¹⁹⁰ In this regard, "educational benefits" refers to the intangibles underrepresented minorities bring to the classroom setting. Thus, when the Court identified racial diversity as a crucial component of any diversity project, its rationale did not escape this basic logic.¹⁹¹ Limited in this sense, the Court embraced the concept of a critical mass only as a method for ensuring the attainment of educational benefits. Had the attainment of a critical mass been unrelated to such benefits, Michigan's policy would have amounted to nothing more than "outright racial balancing," a "patently unconstitutional" policy.¹⁹²

187. See *Parents Involved v. Seattle Sch. Dist. No. 1*, 551 U.S. 701 (2007) finding a lack of narrow tailoring in public school integration plan).

188. The "fit" nature of strict scrutiny analysis necessitates comprehensive articulation of the compelling interests underlying a policy. This is especially true in the context of race-consciousness since the Court is particularly demanding of a tight nexus between means (narrow tailoring) and ends (compelling interest).

189. See *Grutter v. Bollinger*, 539 U.S. 306, 328 (2003).

190. *Id.* at 330 (emphasizing that diversity serves the laudable goals of producing "cross-racial understanding," creating livelier classroom discussion, and "prepar[ing] students for an increasingly diverse workforce and society").

191. *Grutter*, 539 U.S. at 329 (justifying the goal of a critical mass in terms of the "educational benefits" that attach to such an end). For an example of the diversity rationale embodying at least part of the "critical mass" side of the equation, see Expert Report of Kent D. Syverud, *Grutter v. Bollinger*, 137 F. Supp. 2d 821 No. 97-75928 (E.D. Mich. 2001) ("[When] a critical mass of underrepresented minority students is present, racial stereotypes lose their force because nonminority students learn there is no 'minority viewpoint,' but rather a variety of viewpoints among minority students.").

192. *Grutter*, 539 U.S. at 329.

The rationale underlying the acceptance of diversity as a compelling interest in *Grutter* provides a useful point of departure for this defense of rescaling. However, focusing exclusively on the “educational benefits” conceptualization of diversity unnecessarily limits the persuasiveness of diversity as a compelling interest. Without discarding the Court’s notion of “educational benefits,” considering the dual motivations underlying rescaling reframes the value of diversity by focusing on the experience of underrepresented students. This reframing occurs by adding modern understandings of the unique harms and challenges arising from token status to traditional explanations for the need of a critical mass. Through this lens, critical mass assumes the additional function of alleviating these burdens through the amelioration of racially hostile environments. In other words, critical mass not only promotes Justice O’Connor’s “educational benefits,” but also mitigates the race-dependent “token-tax” associated with severe underrepresentation.

Rescaling’s role in the evolution of the diversity rationale does not end here. Traditional diversity arguments assume that those students who are denied admission as a result of lower test scores are unqualified.¹⁹³ Under this view, diversity and meritocracy appear mutually exclusive. Thus, the Court perceives diversity policies as antithetical to color-blindness. This understanding of diversity policies explains why Justice O’Connor invokes “educational benefits” as a justification for Michigan’s departure from merit and neutrality.¹⁹⁴ Rescaling rejects this basic premise. While the goal of a critical mass is essential to this policy, rescaling is principally concerned with ensuring fair measures. It is founded upon the recognition that, as a consequence of stereotype threat, the LSAT undermeasures the talent of Black and Latino/a students. This foundational principle disrupts the diversity-meritocracy dichotomy and avoids a zero sum analysis. Properly understood as correcting race-dependent measurement biases, rescaling does not require sacrificing merit. Rather, it attains diversity through the promotion of a more color-blind and meritocratic process.

Situating rescaling within this updated diversity rationale is only the first stage of the compelling interest discussion. As rescaling’s fair measures goals comprise an explicit antidiscrimination project, this analysis must include our second compelling interest: remedying past discrimination.¹⁹⁵

Prior to *Grutter*,¹⁹⁶ remedying past discrimination was the only firmly established

193. See Uma M. Jayakumar, *Can Higher Education Meet the Needs of an Increasingly Diverse and Global Society? Campus Diversity and Cross-Cultural Workforce Competencies*, 78 HARV. EDUC. REV. 1 (2008). For a critique of such arguments, see Luke C. Harris, *Prologue*, 9 MICH. J. RACE & L. 1 (2003); Charles R. Lawrence III, *Two Views of the River: A Critique of the Liberal Defense of Affirmative Action*, 101 COLUM. L. REV. 928 (2001).

194. The Law School articulated this framing. See *Grutter*, 539 U.S. at 315-16 (describing the “non-academic” criteria as assessing an applicant’s “potential ‘to contribute to the learning of those around them’” and the goal of a critical mass to “ensur[e] their ability to make unique contributions to the character of the Law School.”).

195. See *Parents Involved v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 720 (2007); *Adarand Constructors, Inc. v. Peña*, 515 U.S. 200, 237 (1995) (“[G]overnment is not disqualified from acting in response to [past discrimination]”).

196. In *Bakke*, Justice Powell held that diversity was a compelling state interest. However, lower courts struggled to determine the precedential effect of this component of the opinion. *Grutter* definitively estab-

compelling interest in the context of higher education.¹⁹⁷ Justice Powell's opinion in *Bakke* began to shape the Court's narrow conceptualization of remediable discrimination.¹⁹⁸ Central to this narrow lens is the notion that an institution may only engage in race-conscious behavior designed to remedy its own acts of identifiable discrimination.¹⁹⁹ The Court has since adopted the position that the Constitution prohibits many race-conscious policies designed to remedy "societal discrimination."²⁰⁰ To justify this "identifiable"- "societal" divide, the Court emphasized the "problems" that would arise from the "amorphous" and unbound character of societal discrimination.²⁰¹ Societal discrimination troubles the court on multiple levels. Without an identifiable perpetrator, the Court is unable to establish a neutral baseline to guide remedial policies. Justice O'Connor exhibited this anxiety in *Croson*, as she explained that "a generalized assertion that there has been past discrimination in an entire industry provides no guidance for a legislative body to determine the precise scope of the injury it seeks to remedy."²⁰² Unable to determine the victims of "actual" discrimination,²⁰³ there is no way to guarantee that the beneficiaries of race-conscious policies have actually suffered legally cognizable harms. Beneficiaries thus could receive an "undeserved racial preference" at the expense of "innocent" third parties. Framed in this way, the Court views policies designed to remedy societal discrimination or

lished diversity as a compelling interest in higher education. *Regents of Univ. of Cal. v. Bakke*, 438 U.S. 265 (1979). *See also Grutter*, 538 U.S. at 325 ("[C]ourts have struggled to discern whether Justice Powell's diversity rationale . . . is nonetheless binding precedent.").

197. *Bakke*, 438 U.S. at 307 ("The State certainly has a legitimate and substantial interest in ameliorating, or eliminating where feasible, the disabling effects of identified discrimination."). Before *Bakke*, several employment cases established the Court's modern understanding of legally cognizable discrimination. The current standard, first expressly stated in *Washington v. Davis*, 426 U.S. 229 (1976), requires a showing of discriminatory intent. *Davis* expressly limited the disparate impact standard from *Griggs* to Title VII claims of discrimination.

198. In addition to remedying societal discrimination, Justice Powell explicitly denied a compelling interest in (1) "reducing the historic deficit of traditionally disfavored minorities in medical schools and the medical profession," and (2) "increasing the number of physicians who will practice in communities currently underserved." *Bakke*, 438 U.S. at 306.

199. *See id.*

200. *Id.* at 306-07. *See also Richmond v. Croson*, 488 U.S. 469, 496-97 (1989) (citing *Bakke*, 438 U.S. at 307) ("Justice Powell contrasted the 'focused' goal of remedying 'wrongs worked by specific instances of racial discrimination' with 'the remedying of the effects of 'societal discrimination,' an amorphous concept that may be ageless in its reach to the past.'").

201. *See Croson*, 488 U.S. at 497 ("Societal discrimination, without more, is too amorphous a basis for imposing a racially classified remedy . . . [In contrast, t]he State certainly has a legitimate and substantial interest in ameliorating, or eliminating where feasible, the disabling effects of identified discrimination."); *Bakke*, 438 U.S. at 307-10 ("[S]ocietal discrimination [is] an amorphous concept of injury that may be ageless in its reach to the past.").

202. *Croson*, 488 U.S. at 498.

203. There is nothing natural about this particular conception of discrimination. In limiting cognizable discrimination to intent-based acts, the Court is making a normative decision based on the implications of a broader impact-driven framework. *See Washington v. Davis*, 426 U.S. 229, 248 (1976) ("A rule that a statute designed to serve neutral ends is nevertheless invalid . . . would raise serious questions about, and perhaps invalidate, a whole range of tax, welfare, public service, regulatory and licensing statutes"); *see also Paul Brest, In Defense of the Antidiscrimination Principle*, 90 HARV. L. REV. 1 (1976).

racial underrepresentation as nothing more than unconstitutional racial balancing.²⁰⁴

These concerns are inapposite to rescaling, which exemplifies the Court's notion of constitutional race-consciousness.²⁰⁵ Rescaling is immune to the "racial balancing" and "innocent third party" objections. As a fair measure of intervention, rescaling is not designed nor intended to ameliorate the remnants of past or present societal discrimination.²⁰⁶ Rather, it corrects a particular, quantifiable measurement bias present in contemporary admissions policies. Reliance on defective tools forecloses qualified and talented Black and Latino/a students from receiving a race-blind review. The concept of innocent third parties is only intelligible if the beneficiaries of a race-conscious policy are otherwise undeserving. Because rescaling is a corrective mechanism that enables qualified students to display their actual talent, innocent third parties do not exist.²⁰⁷

Instead, in reorienting the cause of underrepresentation from deficient minorities to deficient tests, one could view the rationale underlying rescaling as a reconception of the "innocent third party." The innocents are no longer students failing to obtain admission as a consequence of race-conscious policies. The "harmed innocents" are actually the qualified vulnerable Black and Latino/a students who are denied a mechanism to show the true measure of their talent. Since stereotype threat precludes our current practices from providing a truly race-free, individualized review of student merit, rescaling's effect on the status quo brings us closer to our color-blind ideal. In this sense, rescaling addresses the kind of harm underlying the Court's traditional understanding of legally cognizable and remediable discrimination.

Before proceeding to the narrow tailoring analysis, it is important to briefly recall Justice Powell's forgotten insight. As a means of correcting race-based measurement biases, rescaling is an example of a race-conscious policy that ensures the "fair ap-

204. See *Croson*, 488 U.S. at 498 ("Relief for such an ill-defined wrong could extend until the percentage of public contracts . . . mirrored the percentage of minorities in the population as a whole.").

205. See generally *Croson*, 488 U.S. at 494 ("States and their local subdivisions have many legislative weapons at their disposal both to punish and prevent discrimination and to remove arbitrary barriers to minority advancement.").

206. Rescaling is also distinguishable from Professor Cunningham's "lingering effects" argument, which understands present "patterns of exclusion that match historical group-based discrimination" as a recognized compelling government interest. Clark D. Cunningham, Glenn C. Loury & John D. Skrentny, *Passing Strict Scrutiny: Using Social Science to Design Affirmative Action Programs*, 90 GEO. L.J. 835, 854 (2002). Cunningham's articulation of lingering effects would fall somewhere in between the most amorphous conceptualization of societal discrimination and the specific, quantifiable and locatable discrimination rescaling is meant to remedy. This argument fails to challenge the assumption that "race-blind" criteria such as the LSAT accurately measure student talent. The lingering effects rationale is instead premised on the notion that past discrimination produces inferior opportunities for preparation, conceived as a form of discrimination that justifies race-consciousness.

207. Some may argue that the court's opinion in *Ricci v. Destefano*, 129 S. Ct 2658 (2009), suggests otherwise. While the Court identified expectation rights in that particular factual scenario, it is unclear that such an expectation right could extend to a rescaling policy implemented prior to a testing cycle. Others may argue that due to the variance in stereotype threat—some Black students are affected more than others and some White students are threatened as a result of class—rescaling will create innocent third parties if only some threats are countered and if others are overcompensated. For a response to this critique, see *infra* Part III.C.2.

praisal of each individual's academic promise in light of some cultural bias in grading or testing procedures."²⁰⁸ In recognizing the non-preferential nature of such a policy, Justice Powell foreshadowed the potential for rescaling to reshape our traditional understanding of the compelling interests of diversity and remedying discrimination.

C. *Fatal in Fact? Narrow Tailoring*

At this point, we have situated rescaling within two, interdependent compelling interests. It is now necessary to demonstrate that rescaling is narrowly tailored to achieve these ends. Narrow tailoring includes several components:²⁰⁹ (1) the racial classification must not be overinclusive or underinclusive;²¹⁰ (2) the classification must remain flexible and treat people as individuals;²¹¹ (3) the classification should include a sunset;²¹² and (4) there must be no race-neutral alternative.²¹³ I conclude by addressing each component in turn.²¹⁴

1) Policy Must Not Be Overinclusive or Underinclusive

Opponents of rescaling will argue that the policy is overinclusive and underinclusive. These claims will exist even if opponents concede that stereotype threat impacts the performance of vulnerable Black applicants. The objection arises from the recognition that stereotype threat does not impact all Black students equally. Even controlling for vulnerability, a uniform rescaling fails to account for individual differences in stereotype threat susceptibility. Thus, a four-point rescaling will under-compensate certain students, and over-compensate others.²¹⁵

Ultimately, this is an objection to rescaling's lack of absolute precision. While this is a valid critique, we should not overstate the likely over/under-inclusiveness of this policy. Informed by our best understandings of stereotype threat, rescaling is designed to mitigate overcompensation.²¹⁶ The presence of a 150-point floor, below

208. *Regents of Univ. of Cal. v. Bakke*, 438 U.S. 265, 306 n.43 (1979).

209. This operationalization of narrow tailoring was adopted from Ian Ayres & Sydney Foster, *Don't Tell, Don't Ask: Narrow Tailoring after Grutter and Gratz*, 85 TEX. L. REV. 517 (2007).

210. *Croson*, 488 U.S. at 469 (invalidating a government set-aside because the policy included racial and ethnic groups who could not possibly have been the victim of previous discrimination).

211. In the educational context, policies may not involve "strict quotas" or hard factors that make race the determinative factor in an applicant's review. *See, e.g.*, *Grutter v. Bollinger*, 539 U.S. 306, 334 (2003) ("Truly individualized consideration demands that race be used in a flexible, nonmechanical way . . . To be narrowly tailored, a race-conscious admissions program cannot use a quota system.").

212. *See, e.g.*, *Grutter*, 539 U.S. at 342 ("Race-conscious admissions policies must be limited in time."); *Adarand Constructors, Inc. v. Peña*, 515 U.S. 200, 238 (1995); *Croson*, 488 U.S. at 510. For a relevant response to this element of narrow tailoring, see Kang & Banaji, *supra* note 14, at 1116 ("[S]teps taken to provide fairer measures of merit can sunset when we demonstrate that mismeasurement is no longer taking place. The narrow tailoring here is obvious.").

213. *See, e.g.*, *Grutter* 539 U.S. at 339 ("Narrow tailoring does . . . require serious, good faith consideration of workable race-neutral alternatives."); *Croson*, 488 U.S. 510 (condemning the policy in part for failure to attempt race-neutral alternatives).

214. For an argument in favor of applying social science findings to the narrow tailoring analysis of race-conscious programs, see Cunningham, Loury & Skrentny, *supra* note 206.

215. This argument assumes a large deviation in effect sizes. It is possible that stereotype threat similarly impacts vulnerable students and produces a relatively small standard deviation.

216. For constitutional purposes, this is a much larger concern than undercompensation.

which rescaling does not occur, minimizes overinclusiveness by limiting rescaling to those students that stereotype threat is most likely to affect.²¹⁷ Since the science tells us that stereotype threat is most detrimental to the vanguard of the group, restricting rescaling to students scoring one standard deviation above their group mean tracks the boundaries of the phenomenon.²¹⁸

Notwithstanding the existence of a floor, demanding absolute precision is asking for more than the present system currently provides. Standardized test makers refute the notion that individual scores accurately measure student talent. The LSAT is no different. The Law School Admission Council (LSAC), which creates and administers the LSAT, is open about the fact that individual scores are far from precise.²¹⁹ Beyond concerns with predictive validity, an individual's score does not even represent their "true talent," but rather falls inside a range of scores within which exists the individual's actual proficiency.²²⁰ This range is currently represented by a 7-point band, which encompasses the test-taker's actual proficiency only 68% of the time.²²¹ This band and level of proficiency makes sense given the standard deviation falls around 8 points.²²² Thus understood, the LSAT results in broad, grotesque chunking. The LSAC has never presented the LSAT as a precise measure of talent. In many ways, the stereotype threat mean effect is a more precise measure than an individual LSAT score. Considering the LSAT's inherent lack of precision, invalidating a rescaling policy for failure to achieve absolute precision would be a severe instance of sacrificing the better for the perfect.

2) Policy Must Remain Flexible and Treat People as Individuals

Opponents will argue that rescaling's allocation of hard points forecloses the possibility of individualized review. This argument characterizes rescaling as a repackaged version of Michigan's invalidated admissions policy in *Gratz*.²²³ Common-

217. Some may challenge the 150-point floor as arbitrary, and point to a lack of evidence supporting the claim that stereotype threat begins to adversely affect test-takers who score beyond this point. This argument again trades on fragile demands for precision. The floor derives from the mean LSAT score for Black students plus the standard deviation, thus placing all students in this group in the top 20% of LSAT takers for their group. There is thus an implicit assumption that the top 20% of Black LSAT takers are the vanguard and most vulnerable to stereotype threat.

218. The presence of a floor should additionally relieve concerns that rescaling limits the ability to maintain elite institutions.

219. See LSAC, WHAT IS A SCORE BAND, <http://www.lsac.org/LSACResources/Publications/PDFs/scorebands.pdf> ("The LSAT, like any standardized test, is not a perfect measuring instrument . . . An error-free score, called a true score, could only be obtained from a hypothetical test that contained no measurement error. The standard error of measurement is used to construct score bands, which are used in score reports to quantify the uncertainty inherent in individual test scores.")

220. *Id.*

221. *Id.*

222. See LSAT TECHNICAL REPORT SERIES, *supra* note 20.

223. In contrast to the Law School's holistic review, the Undergraduate College's 20-point allocation proved fatal. Compare *Grutter v. Bollinger*, 539 U.S. 306, 336-37 (2003) ("The Law School engages in a highly individualized, holistic review of each applicant's file, giving serious consideration to all the ways an applicant might contribute to a diverse educational environment."), with *Gratz v. Bollinger*, 539 U.S. 244, 271-72 (2003) (criticizing undergraduate race-conscious admissions scheme for failing to provide meaningful review of each individual file).

alities between the policies exist. The twenty points allotted to underrepresented minorities in *Gratz* arguably mirrors rescaling's uniform race-based point distribution. Emphasizing this similarity and the Court's hostility toward Michigan's policy, rescaling's inflexible allocation of hard points may appear fatal.²²⁴

While superficially appealing, this analogy breaks down. First, it is fundamentally problematic to judge rescaling's constitutionality by reference to its location on the *Gratz-Grutter* spectrum. Both Michigan policies were designed to promote the compelling interest of diversity. Thus, the Court's narrow tailoring analysis occurred with respect to this particular goal.²²⁵ Since rescaling is not limited to diversity interests,²²⁶ judging its constitutionality cannot occur through a simple analogy to *Gratz*. In *Gratz*, the Court's hostility to inflexible points was based on the notion that the point distribution, as a means to promote diversity, contravened meritocratic review.²²⁷ Nothing linked the policy to the promotion of color-blind admissions.²²⁸

Rescaling is fundamentally different. Based on the interdependent goals of fair measures and an updated understanding of diversity, rescaling requires a narrow tailoring analysis that parallels these interests. Proper analysis must be situated within a baseline that recognizes the four-point measurement bias afflicting vulnerable Black applicants. This measurement bias is the precise type of process defect Chief Justice Rehnquist repudiated in *Gratz*.²²⁹ Reframed in this way, rescaling's utilization of hard parts more accurately describes the amelioration of a race-dependent distortion of student talent and promotes the goal of individualized review.²³⁰ *Gratz* proves to be of limited value in this particular context.²³¹

224. See *Gratz*, 539 U.S. at 272 (“The only consideration that accompanies this distribution of points is a factual review of an application to determine whether an individual is a member of one of these minority groups.”).

225. While an independent prong of analysis, narrow tailoring is unintelligible without reference to a compelling interest. This is the core of any “fit” analysis.

226. As mentioned in Part III.B, even with respect to diversity, the rationale underlying rescaling moves beyond the limited notion of diversity Justice O'Connor articulated in *Grutter*.

227. *Gratz*, 539 U.S. at 271 (citing Powell, J) (explaining that a ‘plus’ factor was constitutionally permissible because it may “allow for ‘[t]he file of a particular black applicant [to] be examined for his potential contribution to diversity without the factor being decisive when compared, for example, with that of an applicant identified as an Italian-American if the latter is thought to exhibit qualities far more likely to promote beneficial educational pluralism”).

228. Even when reconceptualizing merit in terms of the ability to enhance diversity, the *Gratz* Court understood the Michigan policy as a denial of individualized review. See *id.* at 273 (“Instead of considering how the differing backgrounds, experiences, and characteristics of students A, B, and C might benefit the University, admissions counselors reviewing LSA applications would simply award both A and B 20 points because their applications indicate that they are African-American.”).

229. *Id.* (condemning a program in which “any single characteristic automatically ensure[s] a specific and identifiable contribution to a university’s diversity”).

230. Even accepting this argument, opponents may argue that the policy still fails to escape strict scrutiny because of the determinative nature of the point allocation. See *id.* at 273 (“[T]he effect of automatically awarding 20 points is that virtually every qualified underrepresented minority applicant is admitted.”). This argument again misses the point that the act of rescaling corrects an equally determinative corruption present in the status quo. In other words, the race of applicants is implicitly determinative—but as a function of exclusion.

231. One final objection trades on the notion of intersectionality. See generally Kimberle Crenshaw, *Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color*, 43 STAN. L. REV.

Additional distinctions between rescaling and the Michigan policy illustrate the superficiality of this analogy. In *Gratz*, Chief Justice Rehnquist evoked Justice Powell's *Bakke* opinion in demanding that admissions systems "consider each applicant as an individual, assess all of the qualities that individual possesses, and in turn, evaluate that individual's ability to contribute to the unique setting of higher education."²³² He summarily invalidated the university's admissions program for failing to live up to such a standard.²³³ Underlying this decision was the concern that awarding twenty points to underrepresented minorities disallowed an "individualized consideration" and that twenty points made race a decisive factor "for virtually every minimally qualified underrepresented minority applicant."²³⁴ Chief Justice Rehnquist's normative and doctrinal conceptualization of an individualized, color-blind process mirrors the world rescaling produces.

In the abstract, the Chief Justice's demand that a policy provide individualized review and eliminate race as a decisive factor appears obvious and desirable. Such reasoning prompted our immediate hostility to the four hypothetical worlds we visited in Part II.A. Now that we have returned to a discussion on rescaling we must not forget the lesson from that exercise. It revealed the blurriness distinguishing our current admissions system (World Four) from overt discrimination (World One). Thus, understanding rescaling as a manifestation of the unconstitutional behavior we observed in World Four requires obscuring the presence of stereotype threat. When informed by the science, World Four reemerges as a more accurate reflection of the current admissions regime. Having reoriented ourselves, we must ask whether contemporary admissions policies deny students the race-free, individualized review Chief Justice Rehnquist demanded. Since stereotype threat produces a race-dependent measurement bias, the Chief Justice's first concern appears to be implicated. Second, LSAT scores are the embodiment of hard, inflexible points, and a four-point swing will be determinative for the majority of students. The second concern appears implicated as well. It is thus arguable that under the Chief Justice's rationale, the status quo is constitutionally suspect. As our opening Scantron hypothetical prompted us to ask, if the status quo prevents vulnerable Black and Latino/a students from the

1241 (1991) (introducing the concept of intersectionality). Rescaling's process of categorizing students on the basis of a single race arguably obscures the multiple, intersecting identities comprising any individual student. Failing to account for this complexity is potentially problematic given the malleable nature of stereotype threat. See Shih, Pittinsky & Ambady, *supra* note 51. This concern may overstated. It is likely that as a consequence of the rule of hypodescent, see Gotanda, *supra* note 168, at 23-27, multiracial individuals will be societally perceived as non-White. Thus, stereotype threat remains relevant.

232. *Gratz*, 539 U.S. at 271. It is important to note that these comments concerned the policy's lack of narrow tailoring. However, applying these demands to the current admissions regime demonstrates how rescaling corrects for the type of process-oriented discrimination Chief Justice Rehnquist condemned in *Gratz*. Also, because the LSA's policy was framed as a diversity-promoting measure, components of Chief Justice Rehnquist's opinion would not inherently reflect a lack of fit with respect to rescaling.

233. The Equal Protection challenge in *Gratz* focused on the policy of the University of Michigan's College of Literature, Science and the Arts (LSA) to award twenty points to applicants from historically underrepresented groups (Black, Latino/a and American Indian). The Chief Justice's objection to the policy arose from the belief that allocating twenty points based on race contravened the Constitution's demand for meritocratic, individualized admissions standards. For a critique of this framing, see Wise, *supra* note 116.

234. *Gratz*, 539 U.S. at 271-72.

individualized review our Constitution demands, how could the Constitution prohibit an intervention that mitigates this defect?

3) Policy Must Have Some Sort of Sunset

In *Grutter*, Justice O'Connor expressed the view that temporally limitless race-conscious policies offend the core of the Fourteenth Amendment.²³⁵ Understanding race-conscious policies as inherently preferential, Justice O'Connor demanded that every such policy possess a "logical end point." She identified "sunset provisions" as a means to satisfy this command.²³⁶ Based on this precedent, opponents will likely identify the existence of a strict sunset as a prerequisite for any rescaling policy.

The response to sunset demands is straightforward. First, rescaling should include a specific sunset provision. However, the sunset should not be based on arbitrary predictions. Rather, rescaling should subside the instant that empirical studies reveal that stereotype threat no longer produces a measurement bias on the LSAT.²³⁷ While there is nothing wrong with temporal goals, we should base our policy decisions on the most compelling and up-to-date scientific data.²³⁸ Kang and Banaji emphasize this point, arguing that the "question [of when to enforce a sunset] is conceptually easy to answer and depends on why the specific measure was adopted in the first place . . . steps taken to provide fairer measures of merit can sunset when we demonstrate that mismeasurement is no longer taking place."²³⁹ Following this rationale, once we observe the absence of measurement biases on the LSAT, we should cease rescaling scores.

4) Policy Must Have No Race-Neutral Alternative

Are race-neutral alternatives available?²⁴⁰ One obvious possibility emerges. Institutions could easily eliminate the site of measurement bias by foregoing reliance on the

235. See *Grutter*, 539 U.S. at 341-43. This concern arises from a conceptualization of race-conscious behavior as inherently preferential. See *Richmond v. Croson Co.*, 488 U.S. 467, 510 (1989) (plurality opinion) (requiring that racial classifications have a termination point because it "assure[s] all citizens that the deviation from the norm of equal treatment of all racial and ethnic groups is a temporary matter, a measure taken in the service of the goal of equality itself"). Since rescaling promotes a more meritocratic system, the anxiety regarding a lack of sunset may be unwarranted. The more appropriate concern may be whether observable and quantifiable measurement defects compel intervention.

236. Justice O'Connor predicted that in twenty five years the Law School's policy would no longer be necessary to further the interest of student body diversity. See *Grutter*, 539 U.S. at 342-43.

237. For a repudiation of the notion that sunsets must be established temporally without reference to actual societal changes, see Kang and Banaji, *supra* note 13, at 1116.

238. See generally Linda Hamilton Krieger, *The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity*, 47 STAN. L. REV. 1161 (1995) (arguing that advances in the understanding of social cognition theory provide compelling evidence to move away from an intent-based antidiscrimination jurisprudence).

239. Kang & Banaji, *supra* note 13, at 1116.

240. Even if stereotype threat demands a race-conscious solution, opponents may challenge the decision to focus solely on Black and Latino/a students. If stereotype threat depresses the performance of White males in high level math programs, why not extend rescaling to such a setting? If our interest is correcting a race-dependent measurement bias, shouldn't we correct it everywhere it exists? The response to this tracks earlier arguments. Rescaling's commitment to fair measures cannot be understood without its companion concern of

LSAT. As discussed above, complete elimination of the LSAT is doubtful.²⁴¹ Considering the deference traditionally afforded educational institutions,²⁴² the Court would not likely demand such a complete renovation of the admissions process. Assuming elimination of the LSAT is not required, are other race-neutral possibilities available?

Opponents may argue that focusing on socio-economic status (SES) in lieu of race provides a suitable alternative.²⁴³ Since stereotype threat has the potential to impair the performance of low-SES individuals, it follows that rescaling on the basis of SES promotes fair measures without resorting to racial classifications. An institution's internal policy considerations may also question why we should rescale for a high SES Black male student yet not for a low SES White female student. If correcting for measurement bias is the actual goal of rescaling, why not remedy this problem through a race-neutral category such as SES?

This argument has merit. There exists no inherent reason not to address measurement biases associated with SES. However, focusing solely on SES fails to remedy race-dependent measurement biases, the particular concern underlying rescaling. Since race-dependent mis-measures exist independently of social status, a policy targeting SES will fail to reach a significant portion of Black and Latino/a students. Further, stereotype threat's detrimental effects are greatest when a particular domain implicates a negative stereotype associated with racial identities. Thus, relying on a measure for SES or gender will undercompensate individuals suffering from a race-dependent threat.²⁴⁴

Contrary to its purpose, the race-neutral alternative proves more over-inclusive and under-inclusive than the original race-conscious plan. This result touches on the inevitable limitations associated with utilizing race-neutral remedies to address race-dependent problems.²⁴⁵ For rescaling, a policy designed to ameliorate a quantifiable defect, it makes little sense to abandon race in lieu of a race-neutral yet less narrowly tailored solution. Professor Ayres echoes this concern, arguing that "when the government has a compelling interest to remedy past discrimination, the narrow tailoring principle should not bar racial classifications that tailor the size of the [correction] to the remedial need."²⁴⁶

racially hostile environments. Since White men do not suffer from this impediment, the necessity for rescaling is limited.

241. See West-Faulcon, *supra* note 29.

242. See Suzanna Sherry, *Foundational Facts and Doctrinal Change*, 2011 U. ILL. L. REV. 145, 156 ("[T]his is not your father's strict scrutiny . . . a cursory examination of *Grutter v. Bollinger* . . . reveals glaring doctrinal inconsistencies.").

243. This argument draws upon *Croson*, in which Justice O'Connor suggested that Richmond could have created a program designed to encourage the participation of small firms, since minority businesses would have inevitably fallen into this race-neutral category.

244. *Id.*

245. See generally Ian Ayres, *Narrow Tailoring*, 43 UCLA L. REV. 1781 (1996) (arguing that "race-neutral means" may be "inconsistent with narrow tailoring and may not be a less restrictive alternative than explicit racial classifications").

246. *Id.* at 1784. Additionally, since an SES-driven rescaling will likely lead to the disproportionate increase in White, low-SES students, the race-neutral alternative would fail to address the secondary concern of achieving a critical mass.

CONCLUSION

Have we found a way to resuscitate Justice Powell's forgotten insight and reinsert it into the collective consciousness of admissions offices and equal protection jurisprudence? Stereotype threat appears to be the answer. Our understanding of the measurement biases produced by this phenomenon provides a narrative of highly confident, talented, and qualified Black and Latino/a students denied the neutral and color-blind review our Constitution demands. This injury extends far beyond the moment of admissions decision-making. Due to the disproportionate exclusion of qualified Black and Latino/a students, the few who enter must face unique, race-dependent challenges associated with their token status. The underperformance that often follows functions as an affirmation of the negative stereotypes that fueled stereotype threat in the first place. This vicious cycle is not inevitable. Through a rescaling policy that promotes fair measures through the correction of measurement biases, we can avoid these harms. And properly situated within Justice Powell's rationale, rescaling is immune from constitutional challenge, as it "is no preference at all."