
STEREOTYPE THREAT AND ANTIDISCRIMINATION LAW: AFFIRMATIVE STEPS TO PROMOTE MERITOCRACY AND RACIAL EQUALITY IN EDUCATION

SAM ERMAN & GREGORY M. WALTON*

A new generation of social science research creates new opportunities to increase fairness and reduce racial inequality in education. This research raises important questions for antidiscrimination law.

Over the past twenty years, research conducted around the world has established that for students subject to pervasive negative intellectual stereotypes, such as African American and Latino students (and many other groups, including, in math and science, girls and women), school contexts that call to mind these stereotypes can produce distraction and anxiety that impede school achievement and contribute to racial disparities. This “stereotype threat” is the default in evaluative, challenging academic environments. Hence, common measures of intellectual ability typically

* Sam Erman is an Assistant Professor of Law at the USC Gould Law School. He holds a J.D. (2007) and Ph.D. in American Culture (2010) from the University of Michigan. Gregory M. Walton is an Assistant Professor of Psychology at Stanford University. He holds a Ph.D. in Psychology from Yale University. The authors thank Charles Abernathy, Scott Altman, Scott Bice, Rebecca Brown, Jessica Bulman-Pozen, David Cruz, Richard Fallon, Richard Friedman, Rachel Godsil, Don Herzog, Jerry Kang, Dan Klerman, Alex Lee, Tom Lyon, Ravi Mahesh, Dina Mishra, David Pozen, J.J. Prescott, Richard Primus, Andrea Reynoso, Camille Rich, Stephen Rich, Zach Ritter, Daria Roithmayr, Nora Salem, Matthew Samet, Michael Shapiro, Dan Simon, Chris Whitman, Leslie Wolf, Morgan Watts, Abby Wood, co-clerks, the USC Gould Law School Faculty Workshop, the University of Michigan Law School Lunchtime Faculty Talk Series, the Smithsonian Institution Postdoctoral Fellowship Program, and the U.S. Supreme Court and USC Gould School of Law libraries.

underestimate minority students' potential. But stereotype threat is not inevitable. Brief exercises can reduce its effects, causing lasting improvements in minority student achievement.

As schools respond to stereotype threat, two questions arise: What policies does antidiscrimination law forbid, permit, or require? And what do difficulties answering those questions reveal about ambiguities central to antidiscrimination law? This Article provides several answers. Because policies that aim to reduce stereotype threat rarely require racial classifications, schools may implement them consistent with constitutional and Title VI equal protection. But, where stereotype threat persists in prior academic environments, schools may wish to account for that bias in admissions processes, for instance by deemphasizing threat-infected measurements or by recalibrating stereotyped students' test scores or grades. Such "score corrections," which seek to account for ability obscured by stereotype threat rather than neutralize stereotype threat at the source, will often involve racial classifications that will be legal only if they are narrowly tailored to a compelling governmental interest. In choosing whether and how to implement score corrections, schools will face tradeoffs between cost, effectiveness, and two sources of legal risk key to equal protection. The first source of legal risk involves whether merit is a compelling governmental interest. Merit stalks equal protection as an ill-defined shadow interest, often treated as compelling yet never so declared. Score corrections would bring this ambiguity to the fore, pushing the Supreme Court to define merit and specify whether it is a compelling interest. The second source of legal risk centers on the Court's insistence that equal protection disfavors visible racial classifications because they elevate groups above individuals. Score corrections that use visible racial classifications to recognize, highlight, and credit individuals' otherwise latent ability challenge the Court's logic. They present an instance where visible racial classifications promote fair and valid evaluation of individuals.

INTRODUCTION	309
I. THE CHALLENGE	321
A. THE NEED FOR NEW SOLUTIONS	321
B. STEREOTYPE THREAT AND ITS REDUCTION	323
1. Stereotype Threat.....	324
2. Threat Reduction	329
II. IDEAL AFFIRMATIVE MERITOCRACY	339

A. EQUAL PROTECTION VALUES AND JUSTICE KENNEDY'S ASPIRATION	340
B. EQUAL PROTECTION APPLIED.....	343
C. DISPARATE EFFECTS	344
III. ACCOUNTING FOR BIAS IN PAST MEASURES.....	346
A. OVERVIEW.....	347
B. STRICT SCRUTINY	350
1. A Compelling Interest in Merit?.....	351
2. Narrowly Tailored to What End?	362
C. DISPARATE EFFECTS	371
CONCLUSION.....	377

INTRODUCTION

Given the persistence of large group differences in academic achievement, jurists, lawmakers, and the public frequently view efforts to promote racial equality and efforts to promote merit in education as irreconcilable. Their resulting ambivalence has condemned antidiscrimination law to incoherence and many egalitarian policies to stagnation and retreat. Drawing on a new generation of social science research, this Article argues that the conflict is more apparent than real. Schools can promote racial equality by pursuing merit. Many policies to do so would raise no equal protection problems. They are a legal, low-cost, and immediately available way for schools to reduce inequality meritocratically. Other such policies do raise equal protection problems. These provide jurists new points of entry into tangles at the core of antidiscrimination doctrine.

Federal authorities and the broader public value racial equality, but worry that its pursuit comes at the expense of merit. Majorities of Supreme Court Justices perceive a national obligation to reduce racial inequality and a likelihood that attempts to do so will not be meritocratic.¹ Recent polls find that large majorities of Americans consider race to be a serious, inadequately addressed problem and yet oppose racial “preferences.”²

1. See, for example, the splintered opinions in *Parents Involved in Community Schools v. Seattle School District No. 1*, 551 U.S. 701 (2007).

2. Compare Sam Roberts, *Racial Equality Is Still a Work in Progress, Survey Finds*, N.Y. TIMES (Aug. 22, 2013), <http://www.nytimes.com/2013/08/23/us/americans-see-racial-equality-as-a-work-in-progress-pew-poll-finds.html> (“Fully 80 percent of all Americans say at least some more needs to be done.”), with Scott Clement, *Poll: Most Oppose Race-Based College Admissions Programs*, WASH. POST (June 12, 2013), <http://www.washingtonpost.com/politics/2013/06/11/4aee6cf8-d2b9->

Federal legislation and regulation posit a similar conflict. Both bar many practices that produce disparate racial impacts while generally exempting policies necessary to select the candidates most likely to perform best.³ In the employment context, which is not the focus of this Article, statutory provisions both bar adjustments on the basis of race to employment exam scores⁴ and permit reliance on bona fide employment exams.⁵ Regulations governing most schools carve out an even larger safe harbor, permitting disparate racial outcomes whenever a substantial, legitimate justification exists.⁶

The perception that merit and racial equality are competing commitments has contributed to a body of antidiscrimination law and policy at war with itself under uncertain rules of engagement. The Court has declared landmark federal legislation to combat and remedy race discrimination violative of equal protection.⁷ Having announced that equal protection aims to “eliminate . . . all vestiges of state-imposed segregation,” the Court construes the doctrine as permitting schools to use affirmative action in pursuit of diversity, but not as a means of dismantling most legacies of race discrimination.⁸ The Court’s equal protection jurisprudence also leaves unclear what acts trigger “strict scrutiny,”⁹ what goals can allow acts to survive that review,¹⁰ and what relationship an act and goal must share to survive.¹¹ Today, recipients of federal funds find themselves

11e2-8cbe-1bcbee06f8f8_story.html (“Three quarters of Americans . . . oppose allowing universities to consider race when selecting students . . . ”).

3. See 42 U.S.C. § 2000e-2(k) (2012) (Title VII); *infra* note 150 and accompanying text (Title VI).

4. See 42 U.S.C. § 2000e-2(l) (forbidding employers “to adjust the scores of, use different cutoff scores for, or otherwise alter the results of, employment related tests on the basis of race”).

5. See 42 U.S.C. § 2000e-2(h) (permitting “different standards of compensation . . . pursuant to a bona fide . . . merit system”); *infra* notes 222, 224.

6. See *infra* note 152 and accompanying text.

7. See *Shelby Cnty., Ala. v. Holder*, 133 S. Ct. 2612, 2631 (2013) (striking down the coverage formula of the Voting Rights Act).

8. *Swann v. Charlotte-Mecklenburg Bd. of Educ.*, 402 U.S. 1, 15 (1971). See also *Grutter v. Bollinger*, 539 U.S. 306, 331–34 (2003) (upholding a law school admissions affirmative-action program that aimed to produce diversity); *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 498, 511 (1989) (finding an equal protection violation where a city, based on “a generalized assertion that there has been past discrimination in an entire industry,” pursued remedial action by awarding contracts in that industry in part on the basis of race).

9. E.g., Reva B. Siegel, *The Supreme Court 2012 Term—Foreword: Equality Divided*, 127 HARV. L. REV. 1, 48–49 (2013); Stephen M. Rich, *Inferred Classifications*, 99 VA. L. REV. 1525 (2013).

10. See *infra* Part III.B.1 (discussing meritocracy as a possible compelling interest to justify score corrections).

11. See *infra* Part III.B.2 (discussing how race-based score corrections might be narrowly tailored to advance the interest of meritocracy, and what “narrowly tailored” might mean in this

caught between regulatory demands that they eliminate many racial disparities and judicial intimations that the Constitution may render doing so illegal.¹² On the policy front, as jurists, lawmakers, and citizens have increasingly come to perceive equality and meritocracy as irreconcilable ideals, many egalitarian efforts have stagnated and gone into retreat.

Contrary to official and popular perception, equality and merit can reinforce one another. A new generation of social science research shows that educational policies can promote racial equality by pursuing merit. It does so by identifying one important cause of racial inequality in education—the psychological climates that characterize common education settings. Members of non-Asian racial minority groups typically perform in school while aware of negative stereotypes that impugn the ability of their group. This awareness can trigger a phenomenon known as *stereotype threat*, which can undermine academic performance and contribute to racial inequality.¹³ Organizations may reduce this threat in school contexts, promoting better performance and reducing inequality. Or, they may correct for its effects post-hoc by, for example, deemphasizing threat-infected measures or recalibrating stereotyped students' test scores or grades. We term such practices, which respectively release and recognize ability that is otherwise latent (i.e., typically obscured by stereotype threat), “affirmative meritocracy.”¹⁴

At stake in the implications of stereotype threat is the meaning and fate of normative commitments at the heart of the American experiment: the notion that all people should be treated as having equal inherent worth and the promise that merit and hard work will be rewarded. These principles animate the founding national promise that “all men are created

context).

12. See, e.g., *infra* Part III.B–C (exploring whether Title VI regulations could require schools to implement score corrections).

13. The seminal work includes Claude M. Steele & Joshua Aronson, *Stereotype Threat and the Intellectual Test Performance of African Americans*, 69 J. PERSONALITY & SOC. PSYCHOL. 797 (1995); Claude M. Steele, *A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance*, 52 AM. PSYCHOLOGIST 613 (1997); Steven J. Spencer, Claude M. Steele & Diane M. Quinn, *Stereotype Threat and Women's Math Performance*, 35 J. EXPERIMENTAL SOC. PSYCHOL. 4 (1999); Claude M. Steele, Steven J. Spencer & Joshua Aronson, *Contending with Group Image: The Psychology of Stereotype and Social Identity Threat*, in 34 ADVANCES IN EXPERIMENTAL AND SOC. PSYCHOL. 379 (Mark P. Zanna ed., 2002); CLAUDE M. STEELE, WHISTLING VIVALDI: AND OTHER CLUES TO HOW STEREOTYPES AFFECT US (2010).

14. Gregory M. Walton, Steven J. Spencer & Sam Erman, *Affirmative Meritocracy*, 7 SOC. ISSUES & POL'Y REV. 1, 2–3 (2013). See also Jonathan Feingold, *Racing Towards Color-Blindness: Stereotype Threat and the Myth of Meritocracy*, 3 GEO. J.L. & MODERN CRITICAL RACE PERSP. 231, 257 (2011) (arguing that while “diversity and meritocracy appear mutually exclusive,” insights from stereotype-threat research “reject[] this basic premise”).

equal,”¹⁵ which Abraham Lincoln’s “new birth of freedom”¹⁶ and the emergence of “one person, one vote”¹⁷ each partially redeemed. And they underlie the American dream itself. Making good on this ideal has sometimes required rethinking familiar and facially attractive commitments in difficult ways. For example, *Brown v. Board of Education* sparked wrenching desegregation battles by rejecting a potentially appealing formal approach to equality—if schools are materially equal, they’re equal—in recognition that social meanings and psychological climates can render circumstances unequal.¹⁸ The impetus for this Article is the conviction that the nation is fast approaching a similar crossroads.

Stereotype threat is a consequential and well-documented social-cognitive phenomenon in educational settings (among other contexts).¹⁹ It describes how stereotypes that deem members of particular groups less capable can, for members of those groups, produce performance-impeding distraction and anxiety. In taking a test or in performing in class, students who face negative stereotypes like African American and Latino students and, in math and science, women, reasonably worry that other people could view them negatively as a consequence of the stereotype. Students from nonstereotyped groups simply do not have to contend with the same stereotype; they need not worry that they could be viewed as representatives of a less able group. Thus even in the same classroom and even when treated in objectively similar ways the two groups face different situations. It is as if students from stereotyped groups run facing a headwind in school while students from other groups do not.²⁰

Field studies show that people’s performance varies systematically with the degree to which their identity seems at risk in a particular setting. One study with a national sample found that black adults performed worse on a vocabulary quiz administered by white interviewers than on the same quiz administered by black interviewers. The achievement gap was cut in half with a black interviewer.²¹ Likewise, large-scale natural experiments

15. THE DECLARATION OF INDEPENDENCE para. 2 (U.S. 1776).

16. Abraham Lincoln, Address Delivered at the Dedication of the Cemetery at Gettysburg, in 7 THE COLLECTED WORKS OF ABRAHAM LINCOLN 22, 23 (Roy P. Basler ed., 1953).

17. *Gray v. Sanders*, 372 U.S. 368, 381 (1963).

18. *Brown v. Bd. of Educ. of Topeka*, 347 U.S. 483, 492–95 (1954). See also, e.g., GEORGE LEWIS, MASSIVE RESISTANCE: THE WHITE RESPONSE TO THE CIVIL RIGHTS MOVEMENT (2006).

19. For a review of evidence of stereotype threat, see, for example, Brief of Social and Organization Psychologists as Amici Curiae Supporting Respondents at 12–17, *Fisher v. Univ. of Tex. at Austin*, 133 S. Ct. 2411 (2013) (No. 11–345), and Walton, Spencer & Erman, *supra* note 14.

20. Cf. *Griggs v. Duke Power Co.*, 401 U.S. 424, 432 (1970) (deeming Title VII to bar “testing mechanisms that operate as ‘built-in headwinds’ for minority groups” absent adequate justification).

21. Min-Hsiung Huang, *Race of the Interviewer and the Black-White Test Score Gap*, 38 SOC.

show that having a same-sex or same-race teacher can raise women and minority students' achievement.²² Another study found that white women performed 50 percent better on a practice calculus final exam when they were assured that the test was fair for both men and women.²³ A fourth study found that simply moving demographic questions routinely asked of Advanced Placement ("AP") Calculus test-takers to the end of the test instead of immediately before would increase the number of women who receive college credit every year by 4700.²⁴

Because stereotypes are pervasive in U.S. society, merely describing a test as evaluative of an intellectual ability that is subject to a negative stereotype can evoke worries about stereotypes and trigger stereotype threat.²⁵ Such messages are ubiquitous in school.²⁶ The consequence is that stereotype threat is "the norm in academic environments," as a recent study concluded.²⁷ The most rigorous estimates suggest that stereotype threat accounts for a substantial proportion of racial achievement gaps—one third or more of the white-Latino gap on the SAT, for instance.²⁸

Policies to reduce the impact of stereotype threat can substantially reduce racial inequalities in education. A single one-hour exercise, for example, sought to protect minority first-year college students' feeling of social belonging on campus against stereotypes by demonstrating that worries about belonging are common and temporary, not proof that "people like me" do not belong.²⁹ This intervention raised black students' grades and reduced the black-white achievement gap by half over the next three years; it also reduced the cognitive salience of negative racial stereotypes

SCI. RES. 29, 33–36 (2009).

22. Scott E. Carrell, Marianne E. Page & James E. West, *Sex and Science: How Professor Gender Perpetuates the Gender Gap*, 125 Q.J. ECON. 1101, 1104, 1109–13 (2010); Thomas S. Dee, *Teachers, Race, and Student Achievement in a Randomized Experiment*, 86 REV. ECON. & STAT. 195, 198–99 (2004); Douglas S. Massey & Mary J. Fischer, *Stereotype Threat and Academic Performance: New Findings from a Racially Diverse Sample of College Freshmen*, 2 DU BOIS REV. 45, 57–61 (2005).

23. Catherine Good, Joshua Aronson & Jayne Ann Harder, *Problems in the Pipeline: Stereotype Threat and Women's Achievement in High-Level Math Courses*, 29 J. APPLIED DEVELOPMENTAL PSYCHOL. 17, 21–23, 25 (2008).

24. Kelly Danaher & Christian S. Crandall, *Stereotype Threat in Applied Settings Re-Examined*, 38 J. APPLIED SOC. PSYCHOL. 1639, 1644–47 (2008).

25. See Steele & Aronson, *supra* note 13, at 808 ("[M]aking African American participants vulnerable to judgment by negative stereotypes about their group's intellectual ability depressed their standardized test performance relative to White participants, while conditions designed to alleviate this threat, improved their performance . . .").

26. E.g., Steele, *supra* note 13, at 618–19 & n.4.

27. Walton, Spencer & Erman, *supra* note 14, at 8.

28. *Id.* at 11.

29. Gregory M. Walton & Geoffrey L. Cohen, *A Question of Belonging: Race, Social Fit, and Achievement*, 92 J. PERSONALITY & SOC. PSYCHOL. 82, 83, 93 (2007).

for black students.³⁰

Another intervention gives students an opportunity to reflect on personally important values. Such “value-affirmation” exercises help people cope with feelings of threat in school more effectively. This intervention, which involves several fifteen-to-twenty-minute writing exercises administered over the course of the school year, has been shown to raise black and Latino students’ grades over the next two to three years in multiple trials, thereby narrowing achievement gaps.³¹

Critically, these interventions simultaneously reduce racial inequality and promote merit. They reduce the impact of an implicit racial threat in the school climate, freeing students to draw on their full range of faculties as they pursue their studies.

In addition to reducing the adverse effects of stereotype threat on students’ experience and achievement, a school may want to account for the impact of stereotype threat on past measures of performance when making selection decisions. When a student has taken an entrance exam in an environment likely to evoke stereotype threat, a school may wish to account for that threat so as to accurately assess the student’s potential in deciding whether to admit that student. To do so, it might choose to deemphasize the test, something that is already occurring.³² It could also

30. Gregory M. Walton & Geoffrey L. Cohen, *A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students*, 331 SCIENCE 1447, 1447–49 (2011); Gregory M. Walton et al., *Two Brief Interventions to Mitigate a “Chilly Climate” Transform Women’s Experience, Relationships, and Achievement in Engineering*, J. EDUC. PSYCHOL. (forthcoming 2015).

31. Geoffrey L. Cohen et al., *Reducing the Racial Achievement Gap: A Social-Psychological Intervention*, 313 SCIENCE 1307, 1308 (2006) [hereinafter Cohen et al., *Racial Achievement Gap*] (discussing a value affirmation intervention that showed improvement in black students’ grades over one semester); Geoffrey L. Cohen et al., *Recursive Processes in Self-Affirmation: Intervening to Close the Minority Achievement Gap*, 324 SCIENCE 400, 400–01 (2009) [hereinafter Cohen et al., *Recursive Processes*] (follow-up study showing that Black students’ grades improved over a two-year period); David K. Sherman et al., *Deflecting the Trajectory and Changing the Narrative: How Self-Affirmation Affects Academic Performance and Motivation Under Identity Threat*, 104 J. PERSONALITY & SOC. PSYCHOL. 591, 594–95, 597–601 (2013) (similar experiment showing improvement in Latino students’ grades that persisted for up to three years). See also Akira Miyake et al., *Reducing the Gender Achievement Gap in College Science: A Classroom Study of Values Affirmation*, 330 SCIENCE 1234, 1235 (2010) (discussing success of intervention in the context of the gender gap in college science achievement); Judith M. Harackiewicz et al., *Closing the Social Class Achievement Gap for First-Generation Students in Undergraduate Biology*, 106 J. EDUC. PSYCHOL. 375, 377–78, 380–82 (discussing success of intervention in the context of first-generation students in college-level introductory biology courses); Geoffrey L. Cohen & David K. Sherman, *The Psychology of Change: Self-Affirmation and Social Psychological Intervention*, 65 ANN. REV. PSYCHOL. 333, 343–47 (2014) (compiling existing research on the effectiveness of value affirmation interventions in education generally).

32. See, e.g., *Colleges and Universities That Do Not Use SAT/ACT Scores for Admitting*

choose to mitigate the bias in measurement by recalibrating the scores of stereotyped test takers. Public opposition to racial preferences and the relative unfamiliarity with stereotype threat of both the legal community and the public make such “score corrections” controversial. But as understanding of stereotype threat grows, pressure to pursue such approaches may mount.

Although stereotype threat is an important cause of racial inequality, it is far from the only one. Intentional discrimination also curtails opportunities, as do organizational practices known as structural discrimination, which disproportionately impede the progress of members of particular racial groups.³³ Social psychologists have coined the term “implicit bias” to describe how non-conscious attitudes about race can shape people’s actions, sometimes in discriminatory ways.³⁴ Prior discrimination can also be self-compounding, just as wealth, education, and connections often beget each other.³⁵

Because stereotype threat is not the sole cause of racial inequality, reducing and correcting for threat will not eliminate all racial disparities in education. That would require an array of approaches equal to the complexity and depth of the problem, including traditional forms of affirmative action such as outreach and support services and efforts to address causes of inequality that arise from racial disparities in social class such as the quality of schooling available to students of different racial groups. Yet, affirmative meritocracy has the potential to make a substantial contribution.

This Article addresses two audiences: educational and political leaders, and the legal community. For the former we provide two rules of thumb:

(1) Implementing affordable, effective ex-ante policies to reduce stereotype threat will generally make legal sense—in ways inaction does

Substantial Numbers of Students Into Bachelor Degree Programs, FAIRTEST: NAT’L CENTER FOR FAIR & OPEN TESTING (last visited Oct. 9, 2014), <http://www.fairtest.org/university/optional>.

33. Note, “*Trading Action for Access*”: *The Myth of Meritocracy and the Failure to Remedy Structural Discrimination*, 121 HARV. L. REV. 2156, 2159–60 & nn.19–26 (2008) (briefly surveying selected scholarly works on structural discrimination).

34. For a helpful overview, see JERRY KANG, NAT’L CTR. FOR STATE COURTS, IMPLICIT BIAS: A PRIMER FOR COURTS 1–2, 4–5 (2009), available at <http://www.ncsc.org/ibprimer>. But cf. Stephen M. Rich, *Against Prejudice*, 80 GEO. WASH. L. REV. 1, 34–44 (2011) (cautioning against making implicit bias the primary measure of discrimination).

35. See generally DARIA ROTHMAYR, REPRODUCING RACISM: HOW EVERYDAY CHOICES LOCK IN WHITE ADVANTAGE (2014); M. Corcoran, *Rags to Rags: Poverty and Mobility in the United States*, 21 ANN. REV. SOC. 237 (1995).

not.

(2) Threat-correcting forms of affirmative meritocracy often raise revealing and important legal questions, with more affordable and effective measures facing among the steepest legal challenges.

We base these conclusions on a review of the application to affirmative meritocracy of two federal authorities at the heart of education anti-race-discrimination law: constitutional equal protection guarantees³⁶ and Title VI of the Civil Rights Act of 1964.³⁷ Because these authorities respectively bind state actors and recipients of federal funds, most schools are subject to their common dictates.³⁸ Our review teaches that schools that do not endeavor to mitigate stereotype threat may risk causing illegal discrimination by creating unjustified racial disparities. Those that do mitigate stereotype threat, if they proceed with care, may lower the risk of liability without engendering substantial additional risk. Ambiguities in equal protection jurisprudence, however, make entirely eliminating risk unlikely.

There is reason to be optimistic. It is possible to design policies that prevent stereotype threat from arising in education contexts and thus improve outcomes among racial minority students. If implemented widely, such policies could bring substantial benefits and thus help solve one source of what the current Court—and particularly its decisive member on these issues, Justice Anthony Kennedy—describes as the urgent problem of

36. U.S. CONST. amend. XIV, § 1. *See also* *Bolling v. Sharpe*, 347 U.S. 497, 500 (1954); Richard A. Primus, *Bolling Alone*, 104 COLUM. L. REV. 975, 978–81 (2004) (discussing the doctrine of reverse incorporation established by *Bolling* and focusing on the rareness with which federal courts invoke that doctrine in cases involving potential race discrimination against members of minority groups).

37. 42 U.S.C. §§ 2000d–2000d-7 (2012).

38. *See id.* § 2000d (“No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance.”); Civil Rights Cases, 109 U.S. 3 (1883) (establishing the Fourteenth Amendment’s state-action requirement); *infra* note 42 and accompanying text (explaining how the Fourteenth Amendment and Title VI are applied). *Compare* *Grove City Coll. v. Bell*, 465 U.S. 555, 573–74 (1984) (holding in a case featuring a college and a federally funded financial aid program that the “program” that may properly be regulated under Title IX” (which prohibits sex discrimination) was the financial aid program, not the college as a whole), *with* Civil Rights Restoration Act of 1987, secs. 3–6, Pub. L. No. 100-259, 102 Stat. 28, 28–31 (codified as amended at 20 U.S.C. § 1687, 27 U.S.C. § 794(b), 42 U.S.C. §§ 2000d-4a, 6107(4)) (overruling and cabining *Grove City College* by declaring that “the term ‘program’ mean[s] all of the operations of . . . a college, university, or other postsecondary institution, or a public system of higher education” for purposes of a variety of civil-rights statutes, including Title VI), *with* Kiera Feldman, *Sexual Assault at God’s Harvard*, NEW REPUBLIC (Feb. 17, 2014), <http://www.newrepublic.com/article/116623/sexual-assault-patrick-henry-college-gods-harvard> (explaining that Patrick Henry College is not subject to Title IX because it accepts no federal funding).

racial inequality. And it would do so within what the Court casts as the best traditions of U.S. antidiscrimination law.

On the other hand, in a world where stereotype threat is pervasive, such prevention measures are not sufficient to eliminate the consequences of stereotype threat for educational achievement. These measures must be supplemented by corrections for biases caused by threat in past performances. If a university reduces stereotype threat among its current students but cannot eliminate a bias on the SAT that affects its applicants, it will find that black students, if accepted, outperform white students with the same SAT scores. Meta-analyses of real-world threat-reducing strategies have found just this.³⁹ In this case, the black students are like runners who ran to a tie when they faced a headwind during a time trial but now win handily when unencumbered. If the university cannot correct for the effects of stereotype threat on measures that contribute to admissions decisions, it will be unable to select the students likely to perform best and will thus not realize the full benefits of having reduced stereotype threat within its environment. This would undermine the incentives a school has to reduce threat among matriculated students. Moreover, non-Asian minority applicants would regularly be denied admission in favor of white and Asian applicants whom they would predictably outperform.

For the legal audience—judges, advocates, and scholars of law—we argue that affirmative meritocracy presents revealing and complex doctrinal questions. One reason is that from the perspective of antidiscrimination law, stereotype threat is counterintuitive. The paradigmatic doctrinal case of racial discrimination involves intentional harm along racial lines motivated by racial animus.⁴⁰ The perpetrator acts upon a bad mental state to directly harm the victim. The victim's mental state is irrelevant to the process. By contrast, stereotype-threat research instructs that even routine and potentially well-intentioned acts like soliciting demographic information or describing a test as evaluative of ability can raise the salience of race in ways that cause members of stereotyped groups to underperform. The person who most immediately triggers such harm may have had unimpeachable motives. The mental processes that most immediately contribute to the injury are not those of

39. 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, 1137 (2009).

40. See, e.g., Richard A. Primus, *Equal Protection and Disparate Impact: Round Three*, 117 HARV. L. REV. 493, 502–05 (2003); *infra* Part II.B (contending that race-based policies to combat stereotype threat lack a discriminatory intent).

any perpetrators⁴¹ but those of the victims.

An example illustrates the legal problems that affirmative meritocracy brings to the fore. Consider a school that seeks to use racial classifications to correct for effects of stereotype threat and thereby increase the predictive value of selection measures; perhaps it recalibrates applicants' test scores to account for stereotype threat. Generally, the Constitution and Title VI would subject the policy to strict scrutiny and uphold it only if narrowly tailored to a compelling governmental interest.⁴² Such a policy might further state interests that the Court has recognized as compelling, such as to remedy the school's own prior discriminatory acts or to promote academic diversity.⁴³ But this would not be the aim of the policy. Rather, the reason for adopting the policy would be to recognize individual applicants' merit. The Court has never defined the concept of merit in the equal protection context nor determined whether it is a compelling interest; it has, however, implied that merit is a value protected by equal protection.⁴⁴

If merit were a compelling interest, then new questions would arise about narrow tailoring.⁴⁵ The Court portrays narrow tailoring as necessary in the context of benign racial classifications for six reasons. One set of these reasons includes the asserted tendency of narrow tailoring to reduce three negative consequences of racial classifications: exacerbation of racial

41. There are perpetrators, of course. Stereotype threat exists because individuals hold and express negative stereotypes. But the practice is so widespread and so attenuated from the countless instances of harm that it causes that any attempt to link bad mental state to specific injury would be quixotic.

42. See *Fisher v. Univ. of Texas at Austin*, 133 S. Ct. 2411, 2419 (2013) (racial classifications "are constitutional only if they are narrowly tailored to further compelling governmental interests" (quoting *Grutter v. Bollinger*, 539 U.S. 306, 326 (2003))); *Alexander v. Sandoval*, 532 U.S. 275, 280–81 (2001) (Title VI proscribes "only those racial classifications that would violate the Equal Protection Clause").

43. See *Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 720 (2007) ("[O]ur prior cases . . . have recognized . . . the compelling interest of remedying the effects of past intentional discrimination." (citing *Freeman v. Pitts*, 503 U.S. 467, 494 (1992)); *Grutter*, 539 U.S. at 325 ("[S]tudent body diversity is a compelling state interest that can justify the use of race in university admissions."). *But cf.* *Palmore v. Sidoti*, 466 U.S. 429, 433 (1984) ("The State, of course, has a duty of the highest order to protect the interests of minor children").

44. See *infra* Part III.B.

45. See, e.g., Ian Ayres & Sydney Foster, *Don't Tell, Don't Ask: Narrow Tailoring After Grutter and Gratz*, 85 TEX. L. REV. 517, 541–65 (2007) (discussing questions raised about narrow tailoring after *Grutter* and suggesting ways to clarify the doctrine); Richard H. Fallon, Jr., *Strict Judicial Scrutiny*, 54 UCLA L. REV. 1267, 1326–32 (2007) (arguing that the Court's narrow tailoring inquiry "contains significant, unresolved ambiguities of which [it] appears startlingly unaware"); *id.* at 1271–73, 1296, 1307–08, 1312; Jed Rubenfeld, *Affirmative Action*, 107 YALE L.J. 427, 436–43 (1997) (discussing the implications of race-based classifications for the strict scrutiny test).

politics, such as white backlash and ethnic patronage; promotion of racial essentialism, or the idea that all members of a racial group are similar; and promotion of racial depreciation, the view that one race is inferior to another.⁴⁶ Stereotype-threat research partly confirms these intuitions by showing how, in some cases, raising the salience of race can cause substantial harm to stereotyped individuals.⁴⁷ The second set of reasons that a majority of Justices provide for narrow tailoring is the tendency to promote dignity, autonomy, and merit by ensuring that people are treated and allowed to act as individuals rather than as mere members of racial groups.⁴⁸ To date, the Court has portrayed these various goals as mutually reinforcing. Thus, in cases in which racial classifications are justified by the goal of diversity, a majority of the Justices have found that individualized selection minimizes racial politics, essentialism, and depreciation; promotes dignity, autonomy, and merit; and maximizes broad-based diversity.

In considering the provocative idea of recalibrating scores to promote merit, however, the reasons that Justices give for requiring narrow tailoring cut in different directions. The question is what the Court means by merit. Justices sometimes equate meritocratic selection with choosing candidates using selection criteria geared toward producing successful performance.⁴⁹ Based on existing research on stereotype threat, that aspiration would be best served by mechanically recalibrating scores along racial lines.⁵⁰ In other instances, however, Justices cast merit as a result of open competition, a vision potentially at odds with mechanical post-hoc score recalibration.⁵¹ Yet research on stereotype threat simultaneously illustrates the fallacy of an open competition in many education contexts.

Other justifications for narrow tailoring also point in different directions when it comes to score corrections. Some may view score corrections as distributing benefits partly on the basis of race and so as threatening to exacerbate racial politics, essentialism, and depreciation while denying those subject to racial classifications the dignity and autonomy that the Court associates with color blindness. Yet these

46. See *infra* Part III.B.2.

47. Steele & Aronson, *supra* note 13.

48. See *infra* Part III.B.2. The Court also sometimes lauds narrow tailoring for smoking out invidious discriminatory intent by testing actions against stated purposes. Because score corrections aim for merit, application of the test would require determining whether score corrections are genuinely meritocratic, a topic taken up *infra* Part III.B.

49. See *infra* notes 188–203 and accompanying text.

50. See *infra* Part III.A.

51. See *infra* notes 212–216.

corrections arise to recognize individual merit. They reveal and account for stereotyped individuals' otherwise overlooked potential by countermanding racially corrosive messages and freeing students to draw on their full range of faculties as they chart their own paths. As such, score corrections counteract racial depreciation and essentialism, promote dignity and autonomy, and render political backlash illegitimate.

Now consider a school that declines to implement low-cost, race-based score corrections and admits white students whom black and Latino students could predictably outperform had they been selected instead. That choice might put the school in violation of federal regulations promulgated pursuant to Title VI, which forbid unjustified disparate racial impacts. The school would then face a dilemma. Should it employ potentially illegal racial classifications, or risk illegal disparate racial impacts?

That equal protection might forbid what disparate impact laws demand also raises a question at the heart of equal protection doctrine. Do disparate-impact laws violate the Constitution?

The Article proceeds in three parts. Part I introduces the problem—inadequate legal and effective means to address racial inequality—and explores a potential solution—reducing stereotype threat.

Part II examines forms of threat reduction that schools may or must implement under the Equal Protection Clause, Title VI, and associated regulations. Because schools will generally be able to reduce stereotype threat without resorting to racial classifications, this part predicts that they will generally be permitted to do so. Moreover, when such policies are sufficiently cost-effective, regulations promulgated pursuant to Title VI may require schools to act.

Although the ideal response to stereotype threat is to seek to mitigate it *ex ante*,⁵² the option may not always be available, as could sometimes be true for selective schools making admissions decisions based upon measures like SAT scores and prior grades. In such cases, schools may seek ways to correct for the effects of stereotype threat after the fact. Part III describes such potential approaches and evaluates their relative utility, cost, and likely legality. The cheapest, most effective approaches may involve racial classifications, which would be subject to strict scrutiny. Whether any of these approaches would be permissible depends on whether merit (or making tests more meritocratic) is a compelling governmental interest and on how the Court resolves ambiguities in narrow tailoring

52. See *infra* notes 157–158 and accompanying text.

jurisprudence. Mechanical race-based score recalibration, which may be the least expensive and most accurate option, is less likely to survive review than costlier, less-predictive individualized selection. But if mechanical recalibration is constitutional, it would likely be required by Title VI regulations.

I. THE CHALLENGE

A broad array of once-promising approaches to addressing racial inequality in education—from desegregation decrees to affirmative action—now appear inadequate. The challenge: to identify new approaches to substantially and legally reduce racial inequality in education. Part I.A reviews the role that the Supreme Court has played in circumscribing policymakers' options. Part I.B presents key insights from stereotype-threat research, including how to reduce stereotype threat.

A. THE NEED FOR NEW SOLUTIONS

Racial inequality in U.S. education is a persistent, pressing problem in need of new solutions. Although there is, of course, enormous within-group variability, average group differences along racial/ethnic lines remain an urgent social problem. According to the National Assessment of Education Progress, commonly called the Nation's Report Card, the average seventeen-year-old who is African-American or Hispanic reads and performs math at the level of the average white thirteen-year-old.⁵³ On the SAT, white students outscore black and Hispanic students by a margin of 225 to 300 points.⁵⁴ Among adults twenty-five to twenty-nine years old in 2012, 40 percent of whites, 23 percent of blacks, and 15 percent of Hispanics had earned a college degree—wider racial gaps than existed in 1990.⁵⁵

Differences in the schools that black and Latino students attend contribute to the disparity. After peaking around the early 1990s, school desegregation has given way to resegregation.⁵⁶ Today, 35 to 45 percent of

53. NAT'L CTR. FOR EDUC. STATISTICS, U.S. DEP'T OF EDUC., THE NATION'S REPORT CARD: NAEP 2012 TRENDS IN ACADEMIC PROGRESS 16–18, 38–40 (2013) [hereinafter 2012 REPORT CARD], available at <http://nces.ed.gov/nationsreportcard/subject/publications/main2012/pdf/2013456.pdf>.

54. See COLL. BD., 2012 COLLEGE-BOUND SENIORS: TOTAL GROUP PROFILE REPORT 3 (2012), available at <http://media.collegeboard.com/digitalServices/pdf/research/TotalGroup-2012.pdf>.

55. SUSAN AUD ET AL., NAT'L CTR. FOR EDUC. STATISTICS, U.S. DEP'T OF EDUC., THE CONDITION OF EDUCATION 2013, at 10 (2013), available at <http://nces.ed.gov/pubs2013/2013037.pdf>.

56. GARY ORFIELD ET AL., THE CIVIL RIGHTS PROJECT, E PLURIBUS . . . SEPARATION: DEEPENING DOUBLE SEGREGATION FOR MORE STUDENTS, 17–19, 22–23, 32–33, 76–77 (2012), available at <http://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/mlk->

black and Latino students attend “intensely segregated schools” that are often “associated with concentrated poverty.”⁵⁷ Such “[r]acially and socioeconomically isolated schools” generally provide students notably inferior educational opportunities and outcomes.⁵⁸

Yet gains have been made. Two generations ago, spurred by the civil rights movement, U.S. institutions launched ambitious efforts to combat racial inequality in education. Federal courts declared segregation in public schools—and elsewhere—illegal and insisted that it be “eliminated root and branch.”⁵⁹ Civil rights organizations deemed official actions resulting in unjustified, disparate racial impacts to be unconstitutional.⁶⁰ Universities launched affirmative-action programs to remedy discrimination and open education to all.⁶¹ As a result, educational opportunity expanded and racial achievement gaps narrowed.⁶²

But today momentum has waned and reversed.⁶³ One reason is success. Partly as a result of statutes that bar it, openly discriminating on the basis of race has become rarer—a situation that federal law now serves to preserve.⁶⁴ Similarly, anti-segregation decisions and legislation have deprived Jim Crow of legal sanction.

Courts have also contributed. Civil rights lawyers now aspire to convince the Supreme Court not that the Constitution bars racial disparities

national/e-pluribus...separation-deepening-double-segregation-for-more-students/orfield_epluribus_revised_complete_2012.pdf.

57. *Id.* at 19, 27.

58. *Id.* at 6.

59. *Green v. Cnty. Sch. Bd. of New Kent Cnty.*, 391 U.S. 430, 438 (1968).

60. See, e.g., Brief of the N.A.A.C.P. Legal Defense and Educational Fund, Inc. as Amicus Curiae, *Washington v. Davis*, 426 U.S. 229 (1976) (No. 74-1492) (advocating affirmance of a lower court decision that a police employment exam that disproportionately and unjustifiably disfavored black applicants was unconstitutional).

61. See, e.g., TERRY H. ANDERSON, THE PURSUIT OF FAIRNESS: A HISTORY OF AFFIRMATIVE ACTION 150–51 (2004).

62. See, e.g., 2012 REPORT CARD, *supra* note 53, at 2 (describing a narrowing of racial achievement gaps in reading and mathematics from 1971 to 2012); Inst. of Educ. Sciences, *Chapter 3: Postsecondary Education*, in *Digest of Education Statistics: 2011*, NAT'L CTR. FOR EDUC. STATISTICS (2012), http://nces.ed.gov/programs/digest/d11/ch_3.asp (reporting that Hispanic, Asian/Pacific Islander, and black enrollment in postsecondary degree-granting institutions together rose from 14 percent of the total in 1976 to 33 percent of the total in 2010).

63. This is not to say that innovation has ceased. See, e.g., No Child Left Behind Act of 2001, 20 U.S.C. §§ 6301–6578 (2012).

64. See, e.g., Title VI, Civil Rights Act of 1964, 42 U.S.C. § 2000d (2012) (programs that receive federal funds); Title VII, Civil Rights Act of 1964, *id.* § 2000e-2(a) (employers). Cf. *Shelby Cnty., Ala. v. Holder*, 133 S. Ct. 2612, 2632 (2013) (Ginsburg, J., dissenting) (“Congress concluded that . . . continuance [of section 5 of the Voting Rights Act] would facilitate completion of the impressive gains thus far made; and . . . guard against backsliding.”).

in outcomes, but that it permits attempts to remedy disparities.⁶⁵ In public education, courts have dissolved desegregation decrees as having eliminated vestiges of prior *de jure* segregation and as inappropriate vehicles for reducing impacts of residential segregation on racial concentrations in school.⁶⁶

Constraints on traditional affirmative action have emerged as judicial decisions, public opinion, and politics reinforce one another. The Court now subjects such policies to exacting scrutiny and finds many wanting.⁶⁷ The U.S. public has also grown skeptical, with large numbers disfavoring express racial preferences.⁶⁸ That reality makes it difficult to implement and sustain such policies. As a result of statewide ballot initiatives in California and Michigan, for example, those states bar considerations of race that federal law permits.⁶⁹

This Article contributes to the search for paths forward by examining forms of affirmative meritocracy the current Court may permit. While judicial embrace will not guarantee success, judicial disapproval would assure failure. Affirmative meritocracy, we argue, provides a promising approach to reducing racial inequality.

B. STEREOTYPE THREAT AND ITS REDUCTION

This section introduces threat-reducing forms of affirmative meritocracy and their potential to reduce racial disparities in education. Brief interventions that greatly reduce racial achievement gaps across periods of years are possible because of two interrelated social-psychological insights. As Part I.B.1 discusses—and as jurists frequently

65. See, e.g., Brief of Amici Curiae Lawyers' Committee for Civil Rights under Law et al., at 6–8, Ricci v. DeStefano, 557 U.S. 557 (2009) (Nos. 07-1428 & 08-328) (“[A]n employer’s attempt to determine whether there is a less discriminatory promotional test which would apply equally to all in the promotional pool” is not a race-based classification that should receive strict scrutiny.). See also Richard Primus, *The Future of Disparate Impact*, 108 MICH. L. REV. 1341, 1343–44 (2010).

66. See, e.g., Freeman v. Pitts, 503 U.S. 467 (1992) (permitting dissolution of desegregation decree); Pasadena City Bd. of Educ. v. Spangler, 427 U.S. 424, 435–37 (1976) (indicating that the goal of school desegregation programs is to end *de jure* segregation, not counteract residential segregation); Swann v. Charlotte-Mecklenburg Bd. of Educ., 402 U.S. 1, 31–32 (1971) (similar).

67. See, e.g., Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265, 271–72 (1978) (higher education); Gratz v. Bollinger, 539 U.S. 244, 274–75 (2003) (same). See also City of Richmond v. J.A. Croson Co., 488 U.S. 469 (1989) (contracting); Adarand Constructors, Inc. v. Pena, 515 U.S. 200 (1995) (same).

68. Clement, *supra* note 2.

69. See CAL. CONST. art. 1, § 31; MICH. CONST. art. 1, § 26. See also Schuette v. Coal. to Defend Affirmative Action, No. 12-682, slip op. at 16–18, (S. Ct. Apr. 22, 2014) (plurality opinion) (reversing a lower court decision that rejected dismissal by a trial court of a challenge to Michigan’s constitutional bar on all considerations of race in public employment, education, or contracting).

overlook—consciousness of negative stereotypes often impedes performance by members of stereotyped groups, a phenomenon known as “stereotype threat.” Part I.B.2 reviews how simple procedures can reduce stereotype threat, identifies key approaches, and lays out their relative costs and effectiveness.

1. Stereotype Threat

Research on stereotype threat began with an intellectual quandary: Studies repeatedly found that African American, Latino, and Native American students with the very same high-school grades and test scores as white and Asian students nonetheless performed markedly worse in college. The racial achievement gap reproduced in college even when students entered with the same qualifications. This effect, termed *underperformance*, is “almost a lawful phenomenon”⁷⁰; yet this portion of the racial achievement gap cannot be accounted for by racial differences in skills or abilities acquired by age eighteen or in the social class factors that contribute to those differences. Moreover, research found the same pattern by gender in math and science: Women with the same high school math scores as men nonetheless “underperformed” in college math and science courses. What could occur in college environments to cause achievement gaps to reproduce? Stereotype threat provides an account: As students move up the academic ladder to take on more challenging material and interact with peers and instructors in less personal ways, they increasingly contend with the risk that others in the academic environment could view them through the lens of a negative stereotype.⁷¹ This fact causes psychological threat, distraction, and anxiety, which undermine academic performance.⁷² As we now know, the very same processes occur earlier in students’ academic experience, although perhaps to a lesser extent. They undermine adolescents’ performance then as they do adults’ performance later.⁷³

A simple experiment illustrates the phenomenon. Black and white college students are divided into two groups. All take a test adapted from the GRE. In one group, researchers introduce the task as a “test” of verbal acumen. Here, blacks do far worse than whites. In the second group, researchers describe the task as a series of puzzles. This time, black and

70. Steele, *supra* note 13, at 615.

71. Walton, Spencer & Erman, *supra* note 14, at 4–5.

72. *Id.*

73. STEELE, *supra* note 13, at 169–84; Steele, Spencer & Aronson, *supra* note 13, at 385–86; Walton & Spencer, *supra* note 39, at 1134, 1137.

white students perform the same (controlling for SAT scores).⁷⁴ As discussed later, research aggregating many such studies shows that, in such circumstances, black students actually perform better than white students with prior performance equated.⁷⁵

Hundreds of experiments demonstrate the stereotype threat effect.⁷⁶ This research has investigated five interrelated questions. When, how, and who experiences stereotype threat? How can it be remedied? And how does stereotype threat affect the meaning of benchmark performance measures, like standardized test scores and grades?

First, research investigating conditions that trigger stereotype threat finds that a wide range of common situations do so, including simple instructions that represent tests as intellectually evaluative and routine demographic queries.⁷⁷

Second, research shows that stereotype threat causes performance decrements by setting in motion diverse deleterious social-cognitive and affective processes, including physiological stress responses, negative thoughts and emotions, efforts to suppress these psychological reactions, and consequent drains on working-memory efficiency.⁷⁸

Third, research finds that any group that faces a negative stereotype can see declines in performance as a result of stereotype threat, including women taking math tests and white men taking tests said to assess why Asians are so skilled in math.⁷⁹ And it is not limited to education. Telling whites that a golf task tests “natural athletic ability” can impede their performance.⁸⁰ Instructions that trigger stereotype threat can undermine memory performance among the elderly.⁸¹ But in education, stereotype

74. Steele & Aronson, *supra* note 13, at 799–801.

75. Walton & Spencer, *supra* note 39.

76. Walton, Spencer & Erman, *supra* note 14, at 5; Hannah-Hanh D. Nguyen & Ann Marie Ryan, *Does Stereotype Threat Affect Test Performance of Minorities and Women? A Meta-Analysis of Experimental Evidence*, 93 J. APPLIED PSYCHOL. 1314 (2008).

77. Spencer, Steele & Quinn, *supra* note 13, at 21–22 (1999); Steele & Aronson, *supra* note 13, at 798–99; Steele, Spencer & Aronson, *supra* note 13, at 379, 389–91.

78. Toni Schmader, Michael Johns & Chad Forbes, *An Integrated Process Model of Stereotype Threat Effects on Performance*, 115 PSYCHOL. REV. 336, 337 (2008).

79. Walton, Spencer & Erman, *supra* note 14, at 5; Spencer, Steele & Quinn, *supra* note 13, at 10–21; Joshua Aronson et al., *When White Men Can’t Do Math: Necessary and Sufficient Factors in Stereotype Threat*, 35 J. EXPERIMENTAL SOC. PSYCHOL. 29, 30, 32–35 (1999).

80. Jeff Stone et al., *Stereotype Threat Effects on Black and White Athletic Performance*, 77 J. PERSONALITY & SOC. PSYCHOL. 1213, 1223–24 (1999).

81. Alison L. Chasteen et al., *How Feelings of Stereotype Threat Influence Older Adults’ Memory Performance*, 31 EXPERIMENTAL AGING RES. 235, 255 (2005); Walton, Spencer & Erman, *supra* note 14, at 5. The phenomenon sometimes occurs in the opposite direction. When members of

threat particularly impedes those to whom widespread, deprecatory intellectual stereotypes attach,⁸² prominently including non-Asian ethnic minorities.⁸³

Fourth, research has investigated remedies for stereotype threat, both on standardized tests and in classroom performance. These interventions, which are described in detail in the next section, can substantially raise minority students' performance.

Finally, research shows that stereotype threat prevents minority students from performing as well as they are capable on benchmark measures of academic performance like grades and test scores; in this sense, minority students' ability is "latent" in many real-world educational situations—not fully reflected by their level of performance.⁸⁴

What is the evidence for this bias? It comes most directly from two meta-analyses, which examine students' performance at different points in time. Because stereotype threat affects individuals most severely when they face challenging material that pushes them to the limit of their abilities, it typically compounds as students move up the education ladder.⁸⁵ Thus, in a group of black and white students with identical prior SAT scores, the black students tend to underperform their white counterparts on subsequent measures like college grades.⁸⁶ But when researchers administer interventions to reduce stereotype threat, they can reverse this dynamic. With intervention not only does black students' performance rise

nonstereotyped groups undertake tasks, they often do so with the knowledge that other groups are stereotyped as inferior. This knowledge can help students perform better, a dynamic known as "stereotype lift." Its effects are robust but tend to be smaller than those of stereotype threat. Gregory M. Walton & Geoffrey L. Cohen, *Stereotype Lift*, 39 J. EXPERIMENTAL SOC. PSYCHOL. 456, 464 (2003). Real-world threat-neutralizing interventions generally do not affect nonstereotyped students and so appear to have no effect on stereotype lift. See, e.g., Gregory M. Walton & Geoffrey L. Cohen, *A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students*, 331 SCIENCE 1447, 1449–50 (2011) (finding that the intervention set African Americans on an upward trajectory); Cohen et al., *Racial Achievement Gap*, *supra* note 31, at 1309 (same); Cohen et al., *Recursive Processes*, *supra* note 31, at 402–03 (same); Sherman et al., *supra* note 31, at 613 (discussing "the relative ineffectiveness of the [intervention] with Whites").

82. See Walton, Spencer & Erman, *supra* note 14, at 5 ("[S]tereotype threat primarily disadvantages people who face pervasive negative intellectual stereotypes . . .").

83. *Id.*

84. Walton & Spencer, *supra* note 39, at 1137.

85. Steele, *supra* note 13, at 617–18.

86. LEONARD RAMIST, CHARLES LEWIS & LAURA McCAMLEY-JENKINS, COLL. BD., STUDENT GROUP DIFFERENCES IN PREDICTING COLLEGE GRADES: SEX, LANGUAGE, AND ETHNIC GROUPS 28–31 (1994), available at <https://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-1993-1-student-group-differences-predicting-college-grades.pdf>; ARTHUR R. JENSEN, BIAS IN MENTAL TESTING (1980); Steele, *supra* note 13, at 620; Walton & Spencer, *supra* note 39 at 1137.

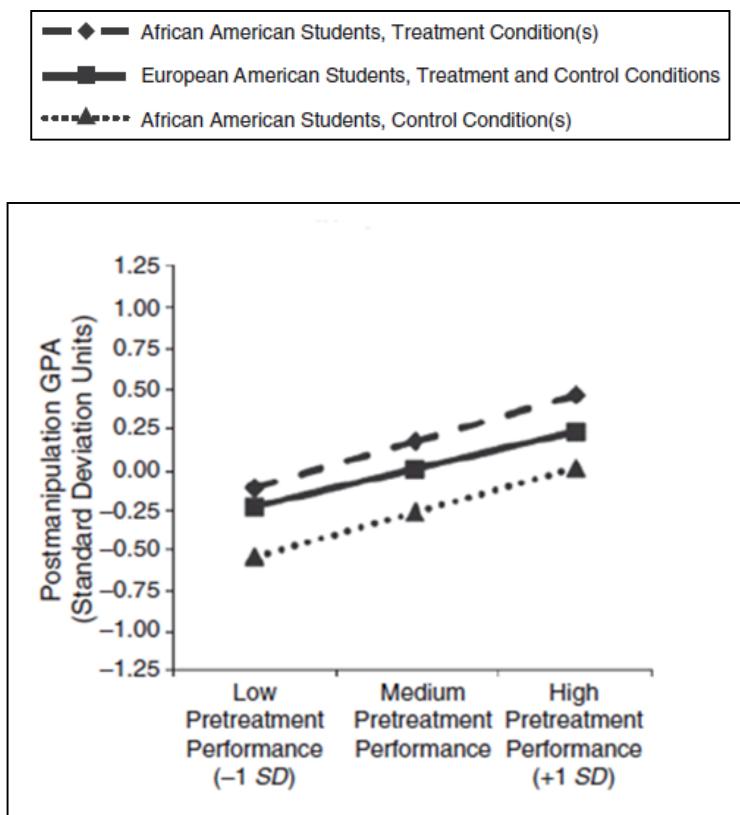
dramatically. Black students also outperform white peers who earned identical prior grades and test scores. The gap between black students in intervention conditions and white students—termed the latent-ability effect—reveals the presence of a bias in the benchmark measures of academic performance.

A recent meta-analysis of results from randomized real-world stereotype-threat interventions confirms this latent-ability effect, as Figure 1 illustrates. Another meta-analysis of laboratory experiments shows the same effect in controlled laboratory settings. These meta-analyses provide evidence that stereotype threat accounts for a quarter of the white-black SAT gap and a third of the white-Latino SAT gap.⁸⁷ Because these estimates reflect only the threat that researchers have identified and neutralized, if anything they likely underestimate the phenomenon.⁸⁸

87. See Walton, Spencer & Erman, *supra* note 14, at 11–12.

88. See *id.* at 9, 11 (finding that if some threat persisted in conditions designed to reduce threat, the ultimate effect sizes underestimate the degree of bias on the benchmark measurement).

FIGURE 1. Classroom Performance of Students in Control Conditions and in Treatment Conditions Designed to Reduce Stereotype Threat⁸⁹



These aspects of stereotype threat provide explanations for two phenomena that have concerned jurists skeptical of race-conscious policies: (1) race gaps that grow as students progress; and (2) difficulties tying racial

89. The figure appears in Walton & Spencer, *supra* note 39, at 1138, which relates: Classroom performance of African American and European American students in control conditions and in treatment conditions designed to reduce stereotype-related threat, as a function of prior performance. The European American control and treatment conditions are combined because they did not yield differences in performance in any study, $t < 1$ GPA = grade point average.

The y-axis plots GPA in standard deviation units because the variance in GPA differed in different studies. The meta-analysis of laboratory experiments included 3180 students in 39 independent samples. The meta-analysis of intervention field experiments included 15,796 students in 3 independent samples. The latent-ability effect—the difference between stereotyped students and nonstereotyped students in “safe” or no-threat conditions—was statistically significant in each meta-analysis at low levels of prior performance ($-1 SD$), at the mean level of prior performance, and at high levels of prior performance ($+1 SD$), all $p < 0.01$.

disparities in exam results to biases in testing content that would render the exams unfair. Richard Sander expresses the first concern: “If it were true that academic indices generally understated the potential of black applicants, then admitted black students would tend to outperform their academic numbers. But[,] . . . [i]f anything, blacks tend to underperform in law school relative to their numbers . . .”⁹⁰ Of course, if stereotype threat pervades law schools, the result and challenged premise would be entirely consistent. Then, one would expect that achievement gaps would persist or widen rather than narrow.⁹¹

Ricci v. DeStefano embodies the second concern. There the Court implied that a test (on which whites disproportionately excelled) was fair because the content appeared to be unbiased. Suggesting that responsibility must thus lie with test-takers, not the test, it quoted an expert’s assertion that “no matter what test the City had administered, it would have revealed a disparity between blacks and whites, Hispanics and whites.”⁹² Stereotype threat deprives the argument of its force. Where contexts produce stereotype threat, unmerited racial gaps can result even when test content is fair.

2. Threat Reduction

As researchers have deepened their understanding of stereotype threat, they have identified ways to reduce it, both in laboratory and in field settings. Even short, relatively inexpensive interventions can have dramatic real-world effects. These interventions are like removing a headwind during a runner’s time trial. They free people to perform to their full potential.⁹³

90. Richard H. Sander, *A Systemic Analysis of Affirmative Action in American Law Schools*, 57 STAN. L. REV. 367, 424 (2004). See also Walton, Spencer & Erman, *supra* note 14, at 6, 25.

91. Compare Fisher v. Univ. of Tex. at Austin, 133 S. Ct. 2411, 2431 (2013) (Thomas, J., concurring) (describing the race gap in high school GPAs and SAT scores among University of Texas at Austin admittees and claiming that “neither the University nor any of the 73 *amici* briefs in support of racial discrimination has presented a shred of evidence that black and Hispanic students are able to close this substantial gap during their time at the University”), with Brief of Experimental Psychologists as Amici Curiae Supporting Respondents at 17, Fisher v. Univ. of Tex. at Austin, 133 S. Ct. 2411 (2013) (No. 11–345) (presenting evidence that after “real-world interventions to reduce stereotype threat,” “stereotyped students perform better than non-stereotyped students who had the same incoming scores”).

92. *Ricci v. DeStefano*, 557 U.S. 557, 572 (2009) (internal quotation marks omitted).

93. For overviews of empirically validated strategies to reduce stereotype threat, see Walton, Spencer & Erman, *supra* note 14; DOROTHY M. STEELE & BECKI COHN-VARGAS, IDENTITY SAFE CLASSROOMS: PLACES TO BELONG AND LEARN (2013); Julio Garcia & Geoffrey L. Cohen, *A Social Psychological Approach to Educational Intervention*, in THE BEHAVIORAL FOUNDATIONS OF POLICY 329 (Eldar Shafir ed., 2013); Cohen & Sherman, *supra* note 31; GREG WALTON ET AL., EMPIRICALLY

Importantly, the specific form of and cues to stereotype threat may differ in different contexts; thus appropriate remedies may differ as well. Social-psychological interventions, including interventions to reduce stereotype threat, are not “magic bullets” or worksheets to be handed out but techniques to remedy maladaptive psychological processes that arise within a particular social context.⁹⁴ To be effective, they must effectively address people’s psychological experience within this context. This may require adapting or tailoring intervention materials and procedures. Below, we review examples of interventions published in the most impactful journals in psychology and science (e.g., *Child Development*, *Psychological Science*, *Journal of Personality and Social Psychology*, *Science*). These interventions illustrate how much student performance can rise through efforts to reduce stereotype threat. They represent starting points for efforts to design effective interventions on large scales within particular organizational contexts.

Buttress Social Belonging. Stereotype threat can cause people to wonder whether they truly belong in school—whether other people will value and include them. Such worries can cause students to interpret negative events on campus like experiences of exclusion or feelings of loneliness as evidence that they do not belong in general. The social-belonging intervention aims to prevent this inference. It gives students a more adaptive narrative for understanding the transition to a new school: namely, that all students worry at first about their belonging but eventually come to feel at home. In the first test of this intervention, first-year students at a selective university participated in a one-hour exercise. Students read survey results in which upper-year students described how they had felt out of place at first but that these feelings declined with time. The students then wrote about how their own experience in the transition to college reflected this same process. Over the next week, students completed daily diaries. During this time, the intervention prevented black students from inferring that they did not belong in general on campus when they encountered everyday adversities. They became more resilient. Across the next three years, their GPAs significantly improved. This effect was mediated by the change in social construal observed in the daily diaries—by the fact that black students no longer saw adversity on campus as proof of non-belonging in general. There was no effect on white students’ grades. In

VALIDATED STRATEGIES TO REDUCE STEREOTYPE THREAT (2012), available at http://www.stanford.edu/~gwalton/home/Welcome_files/StrategiesToReduceStereotypeThreat.pdf.

94. David S. Yeager & Gregory M. Walton, *Social-Psychological Interventions in Education: They’re Not Magic*, 81 REV. EDUC. RES. 267, 274–75 (2011).

total, the intervention caused a 50 percent reduction in the black-white achievement gap from sophomore through senior year.⁹⁵

Although this intervention was originally tested with students already enrolled on campus, large-scale studies have recently adapted it as a prematriculation exercise completed in 30–45 minute online modules by students in the summer before starting college. In this form, the intervention provides students a psychological “roadmap” to common challenges in the transition to college and how to overcome these challenges with time. This intervention has been effective in three trials with full cohorts (90 percent or more) of students at institutions (total N>9500). In a trial among primarily African American, first-generation students exiting an urban charter network, the intervention increased the percentage of students who stayed full-time enrolled in college the next year from 32 to 43 percent. In a second trial at a large, high-quality public university, the intervention increased the proportion of both black and Latino students and of first-generation students who completed the first semester and the first year full-time enrolled, key predictors of on-time graduation, while reducing achievement gaps.⁹⁶ In a third trial at a selective private university, the intervention raised first-year grades among black, Latino, Native American, and first-generation white students, reducing the achievement gap by 35 percent.⁹⁷

Wise Criticism. Critical feedback on academic work can pose a dilemma for students. Does criticism reflect the quality of one’s work to date and a resource for improvement and growth? Or is it a sign that the evaluator dislikes or disrespects you, or even is biased against you? Although this ambiguity can be present for all students, it is most severe for racial-minority students who receive critical feedback from non-minority teachers. If minority students infer that critical feedback results from a teacher’s disrespect or bias, they may disregard it and forego opportunities for growth.⁹⁸ Resolving this ambiguity can help minority students use critical feedback to grow and excel. One study examined the effect of a

95. Walton & Cohen, *supra* note 29, at 94; Walton & Cohen, *supra* note 30 at 1448; Walton, Spencer & Erman, *supra* note 14, at 18.

96. David S. Yeager et al., Brief Psychological Interventions Can Reduce Post-Secondary Achievement Gaps at Scale 6–7 (2014) (unpublished manuscript). See also Paul Tough, *Who Gets to Graduate?*, N.Y. TIMES MAG. (May 15, 2014), <http://www.nytimes.com/2014/05/18/magazine/who-gets-to-graduate.html>.

97. Yeager et al., *supra* note 96, at 7–8.

98. Geoffrey L. Cohen, Claude M. Steele & Lee D. Ross, *The Mentor’s Dilemma: Providing Critical Feedback Across the Racial Divide*, 25 PERSONALITY & SOC. PSYCHOL. BULL. 1302, 1302–03 (1999).

paper-clipped note appended to teachers' feedback on middle school students' essays. When the note read, "I'm giving you these comments so you have feedback on your essay," 17 percent of black students chose to revise and resubmit their essay a week later. But when the note read, "I'm giving you these comments because I have high standards and I know that you can meet them"—disambiguating the reason for the critical feedback—71 percent of black students revised and resubmitted their essay. (Simultaneously, the percentage of white students who revised their essays rose from 62 percent to 87 percent.) When high school students were taught that critical feedback in general reflects high standards and teachers' confidence in students' ability to meet those standards, black students' semester grades rose, reducing the achievement gap by 39 percent. Control-condition black students performed poorly (D or worse) in 43 percent of their courses. In the treatment condition, this dropped to 23 percent.⁹⁹

Growth-Mindset of Intelligence. The belief that intelligence is fixed—that when you encounter academic challenges, it means you have reached the limits of your abilities—can cause students to avoid academic challenges and undermine their tenacity.¹⁰⁰ Teaching students that intelligence can grow with hard work on challenging problems can help all students perform better in school,¹⁰¹ but the benefits can be greatest for students who face negative stereotypes. In one study, mentors encouraged seventh-grade students to adopt a growth mindset of intelligence. At the end of the year, students took the state-mandated math exam. The test revealed small gains from the intervention for boys relative to a control condition but large gains for girls, eliminating a gender gap in math performance.¹⁰²

99. David Scott Yeager et al., *Breaking the Cycle of Mistrust: Wise Interventions to Provide Critical Feedback Across the Racial Divide*, 143 J. EXPERIMENTAL PSYCHOL.: GEN. 804, 808–10, 815–20 (2014); Walton, Spencer & Erman, *supra* note 14, at 19.

100. CAROL S. DWECK, GREGORY M. WALTON & GEOFFREY L. COHEN, ACADEMIC TENACITY: MINDSETS AND SKILLS THAT PROMOTE LONG-TERM LEARNING 5–7 (2011), available at http://web.stanford.edu/~gwalton/home/Welcome_files/DweckWaltonCohen_2014.pdf; Lisa S. Blackwell, Kali H. Trzesniewski & Carol Sorich Dweck, *Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention*, 78 CHILD DEV. 246, 247 (2007).

101. DAVID S. YEAGER ET AL., HOW CAN WE INSTILL PRODUCTIVE MINDSETS AT SCALE? A REVIEW OF THE EVIDENCE AND AN INITIAL R&D AGENDA 6–13 (2013), available at http://www.stanford.edu/~gwalton/home/Publications_files/YeagerPauneskuWaltonDweck%20-%20White%20House%20R%26D%20agenda%20-%205-9-13.pdf; Blackwell, Trzesniewski & Dweck, *supra* note 100, at 258; David Paunesku et al., *Mindset Interventions Are a Scalable Treatment for Academic Underperformance*, PSYCHOL. SCI. (forthcoming).

102. Catherine Good, Joshua Aronson & Michael Inzlicht, *Improving Adolescents' Standardized Test Performance: An Intervention to Reduce the Effects of Stereotype Threat*, 24 APPLIED DEV. PSYCH. 645, 657 (2003); Walton, Spencer & Erman, *supra* note 14, at 19. See also Joshua Aronson, Carrie B.

Value-Affirmation. When people experience stereotype threat, it can seem that all that is relevant is the threat—will other people judge that aspect of your identity and you negatively? In this way, psychological threat is a bit like a physical threat. If you see a tiger in the wild, you won’t call to mind how wonderful your friends are. In school, when students experience stereotype threat, they can lose track of their broader identities and values—those qualities that can make them feel positively about themselves and which can increase their resilience and help them cope with adversity. As a consequence, one way to help students stay engaged and perform better in school in the face of stereotype threat is to give them opportunities to reflect on valued aspects of their identities. “Value-affirmation” interventions aim to do this. In these interventions, students identify values that are most important to them (like relationships with friends and family) and then write for fifteen to twenty minutes as a classroom exercise about why and how these values matter to them. This exercise helps restore students’ sense of self-integrity—their belief that they are good, competent, and moral—and prevents threats like stereotype threat from looming large.¹⁰³

In classic research, seventh-grade students completed either a value-affirmation exercise or a control exercise. The intervention had no effect on white students but it raised black students’ grades, reducing the black-white racial-achievement gap among treatment students by 40 percent. For initially low-performing black students, this intervention (with several additional doses) improved GPA across the final two years of middle school.¹⁰⁴ Similar effects have been observed among Latino adolescents, among whom a value-affirmation intervention raised grades over three years through the transition to ninth grade,¹⁰⁵ as well as among women in physics¹⁰⁶ and first-generation college students in an undergraduate biology course.¹⁰⁷

Fried & Catherine Good, *Reducing the Effects of Stereotype Threat on African American College Students by Shaping Theories of Intelligence*, 38 J. EXPERIMENTAL SOC. PSYCHOL. 113, 123–24 (2002) (“African American [college] students, after just three sessions of advocating the malleability of intelligence, created an enduring and beneficial change in their own attitudes about intelligence.”).

103. Cohen & Sherman, *supra* note 31, at 339–40; Garcia & Cohen, *supra* note 93, at 329; David K. Sherman & Kimberly A. Hartson, *Reconciling Self-Defense with Self-Criticism: Self-Affirmation Theory*, in HANDBOOK OF SELF-ENHANCEMENT AND SELF-PROTECTION 128 (Mark D. Alicke & Constantine Sedikides eds., 2011).

104. Cohen et al., *Recursive Processes*, *supra* note 31, at 400; Walton, Spencer & Erman, *supra* note 14, at 19–20.

105. Sherman et al., *supra* note 31, at 601, 609.

106. Miyake et al., *supra* note 31, at 1236–37.

107. Harackiewicz et al., *supra* note 31, at 385–87.

The Twenty-First Century Program. In an effort to create an institutional program to reduce stereotype threat, the University of Michigan created an honorific, racially diverse dormitory program for first-year students. Among other features, this program included weekly meetings in which students discussed the adjustment to college, learning how this transition was hard for all students (not just them) and ways to overcome common challenges.¹⁰⁸ The program raised black students' first-term GPAs by a third of a point, increased their identification with the university, and reduced the levels of stereotype threat they reported. White students were unaffected.¹⁰⁹

Remove Triggers of Stereotype Threat. Because standardized tests are typically understood as intended to evaluate students' intellectual ability, they generally trigger stereotype threat.¹¹⁰ But small cues can exacerbate this threat. In a foundational laboratory experiment, researchers found that asking black students to indicate their race before a test represented as nonevaluative triggered stereotype threat, undermining scores.¹¹¹ Do demographic triggers increase threat even on tests that are well known to be evaluative? A field experiment of the AP Calculus test found that moving demographic queries from immediately before the test to after the test raised girls' scores; in fact, researchers estimated that, if implemented nationwide, this change would cause 4700 additional girls each year to receive AP Calculus credit.¹¹²

Other cues can, even when inadvertent, trigger stereotype threat, and their absence can mitigate it. For instance, viewing stereotype reinforcing advertisements (e.g., skin products for women) can lower women's aims and outcomes in science and math. The presence of masculine cultural artifacts (e.g., Star Trek posters and video games) in a computer-science

108. Claude M. Steele et al., African American College Achievement: A "Wise" Intervention 10–14 (1998) (unpublished manuscript) (on file with the *Southern California Law Review*). See also Walton & Cohen, *supra* note 29, at 93 (finding that an intervention addressing belonging buffered black students' sense of fit against academic adversity and improved their achievement).

109. Steele, *supra* note 13, at 625–27; Steele et al., *supra* note 108, *passim*; Walton, Spencer & Erman, *supra* note 14, at 20. See also Geoffrey L. Cohen & Claude M. Steele, *A Barrier of Mistrust: How Stereotypes Affect Cross-Race Mentoring*, in IMPROVING ACADEMIC ACHIEVEMENT: IMPACT OF PSYCHOLOGICAL FACTORS ON EDUCATION 303, 305–31 (Joshua Aronson ed., 2002); Walton & Cohen, *supra* note 29, at 94.

110. Walton & Spencer, *supra* note 39, at 1137; Steele & Aronson, *supra* note 13, at 797–99.

111. Steele & Aronson, *supra* note 13, at 805.

112. Danaher & Crandall, *supra* note 24, at 1645–47. But see Lawrence J. Stricker & William C. Ward, *Stereotype Threat in Applied Settings Re-Examined: A Reply*, 38 J. APPLIED SOC. PSYCHOL. 1656, 1661 (2008) ("Inquiring about gender and ethnicity did not affect the test performance of women, Blacks, or other subgroups of students.").

environment can signal a masculine representation of the field and reduce women's sense of belonging and interest in it. Interacting with boorish men during a conversation about engineering can trigger stereotype threat and undermine women's performance.¹¹³ Removing such threat-inducing cues can thus improve outcomes. A broader, important strategy is to reduce the presence in society of negative racial-ethnic stereotypes and prejudice in general; indeed, the degree to which teachers evidence implicit racial bias predicts the size of the racial achievement gap in their classrooms.¹¹⁴

Introduce Cues to Reduce Stereotype Threat. Just as small cues can trigger threat, so too can they reduce it. Presenting students same-race and same-gender role models, for instance, communicates that group members can achieve notwithstanding stereotypes.¹¹⁵ Providing students environments in which minorities are visible and well represented as peers or instructors conveys that members of those groups belong.¹¹⁶ Cues that

113. Paul G. Davies et al., *Consuming Images: How Television Commercials that Elicit Stereotype Threat Can Restrain Women Academically and Professionally*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 1615, 1626 (2002); Sapna Cheryan et al., *Ambient Belonging: How Stereotypical Cues Impact Gender Participation in Computer Science*, 97 J. PERSONALITY & SOC. PSYCHOL. 1045, 1046, 1058 (2009); Walton, Spencer & Erman, *supra* note 14, at 14–15; Christine Logel et al., *Interacting with Sexist Men Triggers Social Identity Threat Among Female Engineers*, 96 J. PERSONALITY & SOC. PSYCHOL. 1089, 1100–01 (2009).

114. See Elizabeth Levy Paluck & Donald P. Green, *Prejudice Reduction: What Works? A Review and Assessment of Research and Practice*, 60 ANN. REV. PSYCHOL. 339 (2009); Thomas F. Pettigrew & Linda R. Tropp, *A Meta-Analytic Test of Intergroup Contact Theory*, 90 J. PERSONALITY & SOC. PSYCHOL. 751 (2006); Jerry Kang & Mahzarin R. Banaji, *Fair Measures: A Behavioral Realist Revision of Affirmative Action*, 94 CALIF. L. REV. 1063, 1071–78 (2006); Linda van den Bergh et al., *The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap*, 47 AM. EDUC. RES. J. 497, 499 (2010).

115. David M. Marx & Phillip Atiba Goff, *Clearing the Air: The Effect of Experimenter Race on Target's Test Performance and Subjective Experience*, 44 BRIT. J. SOC. PSYCHOL. 645, 654 (2005). Cf. David M. Marx & Jasmin S. Roman, *Female Role Models: Protecting Women's Math Test Performance*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 1183, 1191 (2002) (finding that female role models can buffer women's math test performance from the debilitating effects of stereotype threat); Rusty B. McIntyre, René M. Paulson & Charles G. Lord, *Alleviating Women's Mathematics Stereotype Threat Through Salience of Group Achievements*, 39 J. EXPERIMENTAL SOC. PSYCHOL. 83, 88–89 (2003) (finding that women who were reminded about other women's success in math before taking a math test performed better than those who were not); Jane G. Stout et al., *STEMing the Tide: Using Ingroup Experts to Inoculate Women's Self-Concepts in Science, Technology, Engineering, and Mathematics (STEM)*, 100 J. PERSONALITY & SOC. PSYCHOL. 255, 269 (2011) (finding that women exposed to same-sex role models in the STEM fields were more positive about their chosen field and goals than women who had different-sex role models).

116. See Mary C. Murphy, Claude M. Steele & James J. Gross, *Signaling Threat: How Situational Cues Affect Women in Math, Science, and Engineering Settings*, 18 PSYCHOL. SCI. 879, 879 (2007) (finding that seeing a math, science, and engineering environment in which women were outnumbered by men triggered negative physiological responses and feelings of non-belonging among women); Valerie Purdie-Vaughns et al., *Social Identity Contingencies: How Diversity Cues Signal Threat or Safety for African Americans in Mainstream Institutions*, 94 J. PERSONALITY & SOC. PSYCHOL. 615,

signal positive working relationships with majority-group members can reduce worries about being viewed stereotypically among minority-group students and thus stereotype threat; in one natural experiment, having a white roommate at a predominantly white university increased minority students' feelings of social belonging on campus, which led to higher grades.¹¹⁷ Introducing practices that help students feel they are seen as individuals, rather than as tokens of a stigmatized group, can reduce apprehension about negative stereotypes and thus improve performance.¹¹⁸

Help Students Cope with Stereotype Threat. Researchers have identified numerous practices that help students manage the experience of stereotype threat and perform better despite it. These include teaching students about stereotype threat, which helps students attribute anxiety to stereotype threat rather than to a risk of failure;¹¹⁹ teaching students that high arousal can facilitate strong performance, not undermine it,¹²⁰ and

626 (2008) (finding that African American employees in workplace settings where minority representation was low were likely to feel threatened and mistrust their coworkers, but that their trust could be regained if the workplace explicitly articulated valuing diversity). See also 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, 369 (2000) (finding that the underrepresentation of women in a testing environment can cause stereotype threat); Carrell, Page & West, *supra* note 22, at 1139–42 (similar); Dee, *supra* note 22, at 209 (finding that own-race teacher pairings significantly increased the math and reading achievement of minority and nonminority students); Massey & Fischer, *supra* note 22, at 61 (similar).

117. Gregory M. Walton & Priyanka B. Carr, *Social Belonging and the Motivation and Intellectual Achievement of Negatively Stereotyped Students*, in STEREOTYPE THREAT: THEORY, PROCESSES, AND APPLICATION 89, 89–106 (Michael Inzlicht & Toni Schmader eds., 2012); Natalie J. Shook & Russ Clay, *Interracial Roommate Relationships: A Mechanism for Promoting Sense of Belonging at University and Academic Performance*, 48 J. EXPERIMENTAL SOC. PSYCHOL. 1168, 1171 (2012) (showing that having a majority group roommate increased a sense of belonging and academic achievement among minority students); L.J. Aguilar, P.B. Carr & G.M. Walton, *Cues of Working Together Forestall Stereotype Threat* (unpublished manuscript).

118. See Nalini Ambady et al., *Deflecting Negative Self-Relevant Stereotype Activation: The Effects of Individuation*, 40 J. EXPERIMENTAL SOC. PSYCHOL. 401, 403, 406–07 (2004) (finding that individuation—reflecting on one's unique or identifiable characteristics—can protect against stereotype threat); Dana M. Gresky et al., *Effects of Salient Multiple Identities on Women's Performance Under Mathematics Stereotype Threat*, 53 SEX ROLES 703, 711 (2005) (finding that women facing mathematics stereotype threat could improve performance on math tests by reminding themselves of their many social roles and identities).

119. Michael Johns, Tony Schmader & Andy Martens, *Knowing Is Half the Battle: Teaching Stereotype Threat as a Means of Improving Women's Math Performance*, 16 PSYCHOL. SCI. 175, 178 (2005); Walton, Spencer & Erman, *supra* note 14, at 17.

120. Michael Johns, Michael Inzlicht & Tony Schmader, *Stereotype Threat and Executive Resource Depletion: Examining the Influence of Emotion Regulation*, 137 J. EXPERIMENTAL PSYCHOL.: GEN. 691, 699, 701–02 (2008). See also Jeremy P. Jamieson et al., *Turning the Knots in Your Stomach into Bows: Reappraising Arousal Improves Performance on the GRE*, 46 J. EXPERIMENTAL SOC. PSYCHOL. 208, 211 (2010) (finding that reframing arousal as a response to challenges improved performance on math problems).

asking students to contemplate a neutral object (e.g., a red Volkswagen) when they feel anxious, which can prevent students from effortfully suppressing negative thoughts and emotions, and thus raise performance.¹²¹

The promise of interventions to reduce stereotype threat and improve students' performance raises several questions. First, how can schools implement policies to reduce stereotype threat? Because threat arises and can be reduced in many ways that vary by situation, the best way to reduce stereotype threat in a given context is through what we call local empiricism.¹²² Here, a school would review interventions that have proven effective elsewhere, draw on its local knowledge to choose a promising approach, adapt it to its particular context, and then evaluate and learn from the results. Reducing stereotype threat by this method would not be automatic, but it would be achievable.

Second, one might ask whether a school would implement affirmative meritocracy differently depending upon whether its goal is to maximize the performance of all its students, to achieve equal outcomes among students, or to reduce only that threat associated with historically subordinated groups.¹²³

In practice, these various goals would generally lead to similar policies. To the extent that students experience psychological worries that cut across racial lines—like worries about belonging, about the meaning of critical feedback, or about whether intelligence is fixed—then all students benefit from interventions aimed at addressing these concerns.

To the extent that such interventions disproportionately benefit members of non-Asian ethnic-minority groups, there may simply not be another set of available interventions that disproportionately help Asians or whites to a similar degree. These interventions address worries that arise from stereotype threat. If whites and Asians face relatively less stereotype threat, interventions to address this threat may provide them less benefit.

Third, one might inquire whether threat-reducing affirmative meritocracy typically requires racial classifications, which are legally disfavored. It does not. The interventions described above generally treated all participants the same regardless of race; it is thus often possible to reduce stereotype threat without employing racial classifications.

121. Christine Logel et al., *The Perils of Double Consciousness: The Role of Thought Suppression in Stereotype Threat*, 45 J. EXPERIMENTAL SOC. PSYCHOL. 299, 309 (2009).

122. Walton, Spencer & Erman, *supra* note 14, at 25.

123. Rob Reich, *Equality, Adequacy, and K-12 Education*, in EDUCATION, JUSTICE & DEMOCRACY 43 (Danielle Allen & Bob Reich eds., 2013).

Moreover, because stereotype threat can be triggered by contexts that increase the salience of race, threat-reducing interventions that do not categorize by race may often be more effective than interventions that do.¹²⁴ To the extent that interventions that divide people by race do reduce stereotype threat, it will often be feasible to develop race-neutral alternatives.¹²⁵

An additional problem merits mention: whether organizations may account for stereotype threat in contexts from which they select students and how they may do so. The question arises when an organization reduces threat below the levels its applicants have previously faced. Consider a hypothetical university that reduces threat with some success. Assume threat persists on the SAT—and in applicants' high-school environments—for several years. In the interim, the university might desire to admit those students who would perform best if accepted, a group that would include stereotyped students who had somewhat lower SAT scores than their nonstereotyped peers but who would predictably outperform those peers once within a stereotype-safe university. One way to accomplish this goal would be to focus on broader measures of merit, especially indices that are less infected by stereotype threat. Analytic abilities like those assessed by the SAT are perhaps the indicator of merit to which negative intellectual stereotypes apply most directly. Tests that also assess creative and practical forms of intelligence both predict college grades more strongly and yield

124. This is not to say that organizations will succeed by being blind to race. Successful interventions require sophisticated understandings of racial stereotypes and their impacts. Effective approaches to neutralizing stereotype threat may sometimes involve explicit discussions of race. See, e.g., Purdie-Vaughns et al., *supra* note 116, at 626 (finding that African American employees in workplace settings where minority representation was low were likely to feel threatened and mistrust their coworkers, but that their trust could be regained when the workplace emphasized valuing diversity). So long as such discussions occur with all students rather than just those of a particular racial group, no racial classification arises. Notably, although messages that value diversity can be perceived by majority-group students as excluding them, multicultural messages that explicitly encompass both minority and majority groups do not (i.e., "all-inclusive multiculturalism"). Victoria C. Plaut et al., "*What About Me?*" *Perceptions of Exclusion and Whites' Reactions to Multiculturalism*, 101 J. PERSONALITY & SOC. PSYCHOL. 337, 349 (2011); Flannery G. Stevens, Victoria C. Plaut & Jeffrey Sanchez-Burks, *Unlocking the Benefits of Diversity: All-Inclusive Multiculturalism and Positive Organizational Change*, 44 J. APPLIED BEHAV. SCI. 116, 122 (2008).

125. Were a long-term situation to arise in which only racial classifications appeared capable of neutralizing stereotype threat, private groups that both do not accept federal funds and also aim to help students reach college and thrive there could potentially step into the breach. One example of such an organization is the Bright Prospect program. See *Bright Prospect: About Us*, <https://www.brightprospect.org/about/history/> (last visited Jan. 2, 2015); *Prime Award Spending Data*, USASPENDING, <http://www.usaspending.gov/advanced-search> (search for recipient "Bright Prospect") (last visited Jan. 2, 2015) (revealing no federal expenditures on the program); *Sub Award Spending Data*, USASPENDING, <http://www.usaspending.gov/subaward-advanced-search> (last visited Jan. 2, 2015) (same instructions and result).

smaller racial-group differences than traditional college entrance tests like the SAT.¹²⁶ Assessments that include broader measures of preparation for college, including non-cognitive skills like leadership and interpersonal qualities¹²⁷ and self-control and grit,¹²⁸ are also promising. Expanding the measure and definition of merit is likely to reduce, though not eliminate, the effect of stereotype threat on selection decisions. It is also possible to simply eschew metrics that are most infected by threat.¹²⁹ But where such measures are predictive of subsequent performance despite this bias, there may be reasons to retain them. How a school might proceed in this circumstance and what legal, cost, and effectiveness tradeoffs doing so could involve is the topic of Part III.

II. IDEAL AFFIRMATIVE MERITOCRACY

This part argues that the ideal response to stereotype threat—reducing it without resort to racial classifications—is permissible (Part II.B) and may be increasingly required (Part II.C). This likelihood reflects a longstanding duality in the jurisprudence of the decisive Justice on these issues, Justice Kennedy. As Part II.A explains, Kennedy opposes racial classifications while insisting that attempts be made to reduce racial inequality.

126. Robert J. Sternberg et al., *The Rainbow Project: Enhancing the SAT Through Assessments of Analytical, Practical, and Creative Skills*, 34 INTELLIGENCE 321, 344, 347 (2006).

127. Frederick L. Oswald et al., *Developing a Biodata Measure and Situational Judgment Inventory as Predictors of College Student Performance*, 89 J. APPLIED PSYCHOL. 187, 189 (2004).

128. See Angela L. Duckworth & Martin E.P. Seligman, *Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents*, 16 PSYCHOL. SCI. 939, 942 (2005) (“[S]elf-discipline predicted academic performance more robustly than did IQ. Self-discipline also predicted which students would improve their grades over the course of the school year, whereas IQ did not.”). See also Angela L. Duckworth et al., *Grit: Perseverance and Passion for Long-Term Goals*, 92 J. PERSONALITY & SOC. PSYCHOL. 1087, 1098 (2007) (“Across six studies, individual differences in grit accounted for significant incremental variance in success outcomes over and beyond that explained by IQ, to which it was not positively related.”).

129. See Walton, Spencer & Erman, *supra* note 14, at 23 (“[T]his approach may be less appealing when the biased measure is more predictive of outcome measures (and less biased measures are unavailable), and more appealing when the measure is less predictive (and less biased measures are available). When indicators of merit are biased generally—for instance, as a result of bias in common performance contexts—this remedy [of wholly eschewing threat-infected measures] may seem to preclude the use of *any* performance measure in selection decisions.”). Cf. SOC’Y FOR INDUS. & ORGANIZATIONAL PSYCHOLOGY, INC., PRINCIPLES FOR THE VALIDATION AND USE OF PERSONNEL SELECTION PROCEDURES 46 (4th ed. 2003), available at http://www.siop.org/_principles/principles.pdf (“In general, the finding of concern would be evidence of substantial underprediction of performance in the subgroup of interest. Such a finding would generally preclude operational use of the predictor [in selection decisions].”).

A. EQUAL PROTECTION VALUES AND JUSTICE KENNEDY'S ASPIRATION

Justice Kennedy has authored the decisive opinion in every controversial Supreme Court case involving race-conscious measures since Justice Sandra Day O'Connor's 2006 retirement.¹³⁰ As this section elaborates, three contentions animate his approach:

- (1) racial inequality is a large, urgent problem;
- (2) attempts to address racial inequality often counterproductively raise the salience of race; and
- (3) the nation must strive for new, effective approaches that eschew the shortcomings of prior policies.¹³¹

Taken together, these premises cast strict scrutiny as a challenge. Kennedy argues that by closely examining policies with potentially counterproductive effects, judges spur innovation and improvement.¹³²

In *Parents Involved in Community Schools v. Seattle School District No. 1*, Justice Kennedy wrote separately to stress his view that the persistent significance of racial problems in U.S. life demands action. The case involved a challenge to public school student-assignment plans that—in addition to considering student preferences and geography in assigning students into schools—also took into account the racial makeups of schools and applicants. Writing for a plurality of four, Chief Justice John Roberts condemned the racially classificatory policies, asserting: “The way to stop discrimination on the basis of race is to stop discriminating on the basis of race.”¹³³ The government, he contended, should not address racial inequality directly.

Though Justice Kennedy provided the fifth vote invalidating the challenged plans, he disclaimed the plurality’s absolutism. Presuming the salience of race in U.S. life, he wrote: “The enduring hope is that race

130. See *Fisher v. Univ. of Texas at Austin*, 133 S. Ct. 2411 (2013); *Ricci v. DeStefano*, 557 U.S. 557 (2009); *Parents Involved in Cmty. Sch. v. Seattle Sch. Dist.*, 551 U.S. 701, 782 (2007) (Kennedy, J., concurring in part and concurring in the judgment); *Schuette v. Coal. to Defend Affirmative Action*, No. 12-682, slip op. at 1 (S. Ct. Apr. 22, 2014) (plurality opinion); JOAN BISKUPIC, BREAKING IN 210 (2014) (describing *Schuette* as “contentious”). But cf. *Shelby County, Ala. v. Holder*, 133 S. Ct. 2612 (2013) (controversially striking down the coverage formula determining whether a jurisdiction required preclearance under the Voting Rights Act before altering its voting procedures, in an opinion authored by Chief Justice Roberts and joined by Justice Kennedy).

131. See Reva B. Siegel, *From Colorblindness to Antibalkanization: An Emerging Ground of Decision in Race Equality Cases*, 120 YALE L.J. 1278 (2011).

132. See *infra* note 142 and accompanying text.

133. *Parents Involved*, 551 U.S. at 748 (plurality opinion).

should not matter; the reality is that too often it does.”¹³⁴ Acknowledging that problems of racial inequality sometimes require race-conscious and even racially classificatory solutions, he also rejected the plurality’s insistence on unyielding colorblindness as “not sufficient to decide these cases. Fifty years of experience since *Brown v. Board of Education* should teach us that the problem before us defies so easy a solution.”¹³⁵

Turning to the importance of solving race problems, he charged that the “plurality opinion is too dismissive of the legitimate interest government has in ensuring all people have equal opportunity regardless of their race.”¹³⁶ Instead, Kennedy argued, the “Nation has a moral and ethical obligation to fulfill its historic commitment to creating an integrated society that ensures equal opportunity for all of its children.”¹³⁷

Yet, Justice Kennedy did not join the dissenters in voting to uphold the policies. Like the plurality, he worried that even well-intentioned attempts to address pressing social problems can cause great harm, especially when they draw express racial lines.¹³⁸ All race-conscious

134. *Id.* at 787 (Kennedy, J., concurring in part and concurring in the judgment). See also *id.* at 798 (“Due to a variety of factors—some influenced by government, some not—neighborhoods in our communities do not reflect the diversity of our Nation as a whole.”).

135. *Id.* at 788 (citation omitted).

136. *Id.* at 787–88.

137. *Id.* at 797. See also *Schuette v. Coal. to Defend Affirmative Action*, No. 12-682, slip op. at 16 (S. Ct. Apr. 22, 2014) (plurality opinion) (“[A] historical background of race in America that has been a source of tragedy and persisting injustice . . . demands that we continue to learn, to listen, and to remain open to new approaches . . .”).

138. On recent work identifying and exploring this aspect of Justice Kennedy’s jurisprudence, see, for instance, Siegel, *supra* note 131; Michelle Adams, *Is Integration a Discriminatory Purpose?*, 96 IOWA L. REV. 837, 847 (2011) (“Racial classifications are presumptively unconstitutional . . . because such classifications deny individuals their personal rights to be treated with equal dignity and respect, they risk stigmatic harm, and because they may promote notions of racial inferiority and lead to a politics of racial hostility.” (footnotes and internal quotation marks omitted)).

One potential objection to subjecting even transparently virtuous racial classifications to strict scrutiny is that doing so “may impair the coin of the suspect classification doctrine” or “undervalu[e] very strong government interests.” Michael H. Shapiro, *Argument Selection in Constitutional Law: Choosing and Reconstructing Conceptual Systems*, 18 S. CAL. REV. L. & SOC. JUST. 209, 357, 359 (2009). Compare *Bush v. Vera*, 517 U.S. 952, 1010 (1996) (Stevens, J., dissenting) (hypothesizing a government program to protect African American victims of sickle cell anemia, and arguing that “when the state action (i) has neither the intent nor effect of harming any particular group, (ii) is not designed to give effect to irrational prejudices held by its citizens but to break them down, and (iii) uses race as a classification because race is ‘relevant’ to the benign goal of the classification, we need not view the action with the typically fatal skepticism that we have used to strike down the most pernicious forms of state behavior.” (citation omitted)), with *id.* at 984 (plurality opinion) (arguing that such a program would “no doubt” survive strict scrutiny), with Siegel, *supra* note 9, at 63 & n.310 (noting that the Court has repeatedly declined to hear cases regarding whether racial descriptions of at-large suspects constitute invalid racial classifications, an approach that somewhat resembles Shapiro’s avoidance canon). Cf. Peter Westen, *The Empty Idea of Equality*, 95 HARV. L. REV. 537, 539–42,

decision making threatens equal protection values, he wrote, but the “dangers . . . are not as pressing when the same ends are achieved by more indirect means.”¹³⁹

But even as Justice Kennedy joins his colleagues in skepticism of attempts to reduce racial inequality, he insists that “our tradition is to go beyond present achievements, however significant, and to recognize and confront the flaws and injustices that remain. This is especially true when we seek assurance that opportunity is not denied on account of race.”¹⁴⁰ As a result, he urges “bring[ing] to bear the creativity of [all] to find a way to achieve [progress] without resorting to widespread governmental allocation of benefits and burdens on the basis of racial classifications” or, presumably, to policies involving similar equal protection dangers.¹⁴¹

Justice Kennedy insists that in requiring much of policies that aim to reduce racial inequality the Court inspires rather than frustrates their development. As expounded in his dissent in *Grutter v. Bollinger*, unless courts employ “a searching standard to race-based” decisions in order to “force educational institutions to seriously explore race-neutral alternatives,”¹⁴² they “lose the talents and resources of [interested parties] in devising new and fairer ways to ensure individual consideration.”¹⁴³

passim (1982) (rejecting the notion of equality as a distinct basis for any claim, which suggests that those seeking to preserve the “coin” will need to protect it).

139. *Parents Involved*, 551 U.S. at 797 (Kennedy, J., concurring in part and concurring in the judgment). See also *Fisher v. Univ. of Texas at Austin*, 133 S. Ct. 2411, 2418 (2013) (“Distinctions between citizens solely because of their ancestry are by their very nature odious to a free people,’ and therefore ‘are contrary to our traditions and hence constitutionally suspect.’”) (citations omitted) (quoting *Rice v. Cayetano*, 528 U.S. 495, 517 (2000); *Bolling v. Sharpe*, 347 U.S. 497, 499 (1954))); Neil S. Siegel, *Race-Conscious Student Assignment Plans: Balkanization, Integration, and Individualized Consideration*, 56 DUKE L.J. 781, 809–11 (2006) (discussing factors the Court takes into account when considering the net effect on balkanization of using racial criteria, and identifying those factors as including the “symbolic message conveyed by the use of race”; the “burdens that using race imposes on individuals”; “the extent to which consideration of merit is possible”; and “the realization of significant social benefits”).

140. *Parents Involved*, 551 U.S. at 787 (Kennedy, J., concurring in part and concurring in the judgment). See Siegel, *supra* note 9, at 92–93 (Justice Kennedy “seem[s] to understand [his] role in equal protection cases as permitting government to promote diversity and equal opportunity in race-conscious ways, while tightly restricting government interventions of this kind to protect the interests of those who claim the laws are unfair.”). On Kennedy’s idealistic register in *Parents Involved*, see Heather K. Gerken, *Justice Kennedy and the Domains of Equal Protection*, 121 HARV. L. REV. 104, 105 (2007).

141. *Parents Involved*, 551 U.S. at 798 (Kennedy, J., concurring in part and concurring in the judgment).

142. *Grutter v. Bollinger*, 539 U.S. 306, 394 (2003) (Kennedy, J., dissenting).

143. *Id.* at 393. Cf. *Schuette v. Coal. to Defend Affirmative Action*, No. 12-682, slip op. at 16–17 (S. Ct. Apr. 22, 2014) (plurality opinion) (“[A] democracy has the capacity—and the duty—to learn from its past mistakes; to discover and confront persisting biases That process is impeded, not

The danger is that perfection is the enemy of good. Securing policies to promote racial equality is hard work.¹⁴⁴ Success is not foreordained.¹⁴⁵ As the Court raises its standards, many may abandon the fight.¹⁴⁶ Racial inequality could fester. Critics—and much of the U.S. public—contend that it already does.¹⁴⁷

We sound a more optimistic chord. Powerful, inexpensive, and targeted strategies to mitigate stereotype threat have the potential to meaningfully reduce racial inequality while eschewing racial classification and neutralizing racially corrosive messages. This approach may be the white-hatted stranger to antidiscrimination law for whom Justice Kennedy has been waiting.

B. EQUAL PROTECTION APPLIED

We now turn from theory to doctrine. This section contends that schools may implement threat-reducing measures that do not employ racial classifications without inviting suit under the Constitution or Title VI.¹⁴⁸ The Constitution and Title VI subject intentional discrimination—that is, acts taken with racially discriminatory purposes or racial classifications—to strict scrutiny. Only the former trigger is at issue here. Recall that schools will usually be able to implement a set of policies close to their ideal by seeking to maximize students' performances. That aspiration is clearly not a discriminatory purpose. If they follow this course, they will not intentionally discriminate, will not trigger heightened scrutiny, and so will only be subject to rational basis scrutiny. “[D]isparity of treatment” that nonetheless resulted would have a “rational relationship” to the “legitimate governmental purpose” of recognizing and realizing potential

advanced, by court decrees based on the proposition that the public cannot have the requisite repose to discuss certain issues.”).

144. Siegel, *supra* note 131, at 1353.

145. *Id.*

146. *Id.*

147. See, e.g., ERICA FRANKENBERG, CHUNGMEI LEE & GARY ORFIELD, CIVIL RIGHTS PROJECT, HARVARD UNIV., A MULTIRACIAL SOCIETY WITH SEGREGATED SCHOOLS: ARE WE LOSING THE DREAM? 4, *passim* (2003), available at <http://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/a-multiracial-society-with-segregated-schools-are-we-losing-the-dream/frankenberg-multiracial-society-losing-the-dream.pdf> (reporting and criticizing the growing problem of “[r]acial enrollment and segregation in American public schools”); Matthew J. Lindsay, *How Antidiscrimination Law Learned to Live with Racial Inequality*, 75 U. CIN. L. REV. 87, 89 (2006) (“American antidiscrimination law, though unequivocally committed to racial equality in theory, has learned to live with racial inequality in fact.”); Roberts, *supra* note 2 (describing a poll showing that a vast majority of Americans think more needs to be done in pursuit of racial equality).

148. On binding regulations promulgated pursuant to Title VI that create no private right of action, see Alexander v. Sandoval, 532 U.S. 275 (2001); *infra* Sections II.C & III.C.

and so be presumptively legal.¹⁴⁹

The brevity of the argument underscores its importance. Core forms of affirmative meritocracy have the potential to substantially reduce racial inequality. And establishing their legality under equal protection is straightforward.

C. DISPARATE EFFECTS

In fact, many schools may come to be required to implement threat-reducing measures. When recipients of federal dollars decline to implement available strategies to mitigate stereotype threat, they may produce illegal disparate racial effects. Regulations promulgated by federal agencies pursuant to Title VI prohibit recipients of federal funds from employing “criteria or methods of administration which have the effect of subjecting individuals to discrimination because of their race.”¹⁵⁰ Most schools—as well as the Educational Testing Service (“ETS”), which administers the SAT and AP tests—receive federal funds.¹⁵¹ Organizations that decline to implement available strategies will often produce discriminatory effects. A safe harbor exists for practices with a substantial, legitimate justification that cannot be replaced with an equally effective, less-discriminatory alternative.¹⁵² But many practices will not qualify for that protection.

Organizations are only responsible for disparate effects traceable to their criteria and methods. In some cases, research on stereotype threat will provide tools with which to show this causal link. For instance, research might help identify a specific cue that, in a specific context, triggers stereotype threat and produces measurable harm. Examples might include

149. *Armour v. City of Indianapolis*, 132 S. Ct. 2073, 2080 (2012) (quoting *Heller v. Doe*, 509 U.S. 312, 319–20 (1993)). For a rare example of a law that fails rational-basis scrutiny, see *Romer v. Evans*, 517 U.S. 620, 632 (1996) (finding that a law “inexplicable by anything but animus toward the class it affects . . . lacks a rational relationship to legitimate state interests”).

150. E.g., 28 C.F.R. § 42.104(b)(2) (2013).

151. For information on recent federal spending related to ETS, see *Prime Award Spending Data*, USASPENDING, <http://www.usaspending.gov/advanced-search> (enter DUNS Number “002508463,” then click “Search”; then filter by agency (Department of Education) and by fiscal year (2013)) (last visited Jan. 2, 2015) (reporting that in 2013 ETS received more than \$45 million dollars in federal funds from agencies that implemented Title VI regulations). See also 34 C.F.R. § 100.3(b)(2) (prohibiting recipients of funds received under any law administered by the Department of Education to employ “criteria or methods of administration which have the effect of subjecting individuals to discrimination because of their race”); *Agency-Specific Civil Rights Information*, U.S. DEPARTMENT JUST.: C.R. DIVISION, <http://www.justice.gov/crt/about/cor/fedagencies.php> (last visited Jan. 2, 2015) (listing federal agencies that have promulgated regulations pursuant to Title VI and providing links to those regulations).

152. *Title VI Legal Manual—Part VIII.B.: Disparate Impact/Effects*, U.S. DEPARTMENT JUST.: C.R. DIVISION, http://www.justice.gov/crt/grants_statutes/legalman.php#Disparate (Sept. 1998).

asking for demographic information immediately before a standardized test or decorating a study space with successful former students or faculty who happen to all be white. Research might also provide evidence that stereotype threat and measurable associated harm is the result of a school maintaining a campus environment with so few minority group members that they do not form a critical mass.¹⁵³ In other cases, research could reveal that measures of performance upon which a school relies in admissions are infected by threat. Of course, in some instances the harms of stereotype threat may be due to society-wide factors to the exclusion of actions attributable to a given school.

If a practice creates a disparate effect, Title VI regulations require both (1) a substantial, legitimate justification for this effect; and (2) that no equally effective less-discriminatory alternative exists. Under the first requirement, a university may have a substantial, legitimate justification for requesting demographic information: identifying and preventing discrimination and promoting learning by securing a diverse student body. Similarly, the use of test scores in admissions may help schools select high-achieving students necessary to maintaining a reputation for excellence. But not all actions are easily justified. Small, seemingly inconsequential acts can trigger stereotype threat. These practices—like displaying photographs of praiseworthy alumni and happening only to choose those who are white—may lack a substantial, legitimate justification.¹⁵⁴

Under the second requirement, an alternative will be equally effective if it serves the purposes of the organization as well as the original practice without imposing disqualifying additional costs. Schools will often be able to reduce threat through nonclassificatory steps that are equally effective alternatives to doing nothing. This would be true, for instance, of requesting demographic information after rather than before a test, in a preregistration process long before the test instead of the day of, or in

153. See, e.g., Walton et al., *supra* note 30, at 36 (“Although these interventions were successful, it would be far better, if possible, to improve STEM settings themselves . . . [by] increas[ing] the representation of women. Consistent with the concept of critical mass, a community of women may reduce social marginalization and the need for specific intervention.” (citing Henry Etzkowitz et al., *The Paradox of Critical Mass for Women in Science*, 266 SCIENCE 51 (1994))).

154. Of course, this example could raise complex First Amendment concerns. Cf. Eugene Volokh, *How Harassment Law Restricts Free Speech*, 47 RUTGERS L. REV. 563 (1995) (discussing potential conflicts between anti-harassment statutes and free speech values); Benjamin Dower, Note, *The Scylla of Sexual Harassment and the Charybdis of Free Speech: How Public Universities Can Craft Policies to Avoid Liability*, 31 REV. LITIG. 703, 711–20 (2012) (same); Kinglsey R. Browne, *Zero Tolerance for the First Amendment: Title VII’s Regulation of Employee Speech*, 27 OHIO N.U. L. REV. 563, 575 & n.77 (2001) (same and collecting sources). Though beyond the scope of this Article, the topic bears mentioning because many triggers of stereotype threat are communicative acts.

requesting information from schools rather than from individual students—a step that ETS recently decided to take at least with regard to gender on the SAT.¹⁵⁵ It might also be true of many social-belonging and growth-mindset interventions like those discussed in Part I.B.2, which have the potential to be delivered effectively online to large numbers of students at low cost.¹⁵⁶ In other cases, this might not be true. For instance, eschewing threat-infected measures that are predictive of performance—as grades and SAT scores may be—could sacrifice predictive power.¹⁵⁷ The University of Michigan’s Twenty-First-Century Program illustrates an alternative that may not qualify as equally effective for cost reasons. The Twenty-First-Century program reduced certain racial gaps by routing students from a traditional residential program to a special dormitory with unique programming.¹⁵⁸ If this program posed a significant cost above and beyond existing housing arrangements, this expense may render such a program a less-than-equally-effective alternative.

Despite the modest complexities discussed above, this part has a clear bottom line. It makes legal sense to reduce stereotype threat where one can do so without racial classifications.

III. ACCOUNTING FOR BIAS IN PAST MEASURES

In an ideal world, score corrections would be unnecessary. People and organizations would discover and implement ways to reduce stereotype threat throughout society, eliminating the need to account for its presence. But that world is not ours. One goal of this Article is to convince schools to aggressively seek to reduce stereotype threat. The case for doing so is strong, and we are optimistic that increasing numbers of educational

155. Press Release, Coll. Bd., Enhanced SAT Security Measures Announced for 2012–13 Academic Year (Mar. 27, 2012), available at <http://www.cinemnewsnow.com/news/local/Enhanced-SAT-security-measures-announced-for-2012-13-academic-year-144415615.html>.

156. Paunesku et al., *supra* note 101; Yeager et al., *supra* note 96, at 6–7. For discussion, see STEELE & COHN-VARGAS, *supra* note 93; YEAGER ET AL., *supra* note 101; Yeager & Walton, *supra* note 94, at 289.

157. *But cf. supra* note 128 and accompanying text (describing an approach that reduces racial group differences while increasing predictive power).

158. See *supra* notes 108–109 and accompanying text. A special case arises where an intervention that neutralizes threat for stereotyped individuals and thereby substantially increases their performance also eliminates the effects of stereotype lift for nonstereotyped individuals and slightly depresses their performance. Such interventions would be equally effective and less discriminatory so long as the benefits of the unleashed performance outweighed the losses associated with the depressions. While it is possible to imagine scenarios where small losses for many would outweigh large gains by few, such situations would be rare—especially because real world, threat-reducing interventions have generally resulted in no impact on nonstereotyped individuals or modest positive impacts. See Aronson, Fried & Good, *supra* note 102; Yeager et al., *supra* note 99; *supra* note 81.

institutions and testing organizations will undertake the effort. But even in the best case scenario, progress will occur across years, come unevenly, and never be total. Some institutions may reduce stereotype threat effectively. Others may make modest initial gains. Some will delay or decline the project.

Where schools cannot reduce threat—when it arises among applicants who have not yet matriculated, for instance—they may seek to reduce its impact through post-hoc corrections. This part examines such efforts. Part III.A evaluates potential approaches and observes that the cheapest and most effective may use racial classifications. Whether such classifications survive strict scrutiny, Part III.B explains, will depend on how the Court resolves ambiguities concerning compelling governmental interests and narrow tailoring. Part III.C adds that certain score corrections, if otherwise permissible, may sometimes be mandatory under Title VI regulations.

A. OVERVIEW

Correcting for the effects of stereotype threat after the fact rather than reducing it has two serious non-legal drawbacks. It permits stereotype threat—a harmful psychological process that undermines learning as well as performance—to persist.¹⁵⁹ Many forms are also less accurate. When an organization eliminates stereotype threat, it permits each individual to perform unburdened by threat. Corrections that recalibrate scores seek to replicate this result by adjusting how threat-infected measures are used. But as discussed below, this approach will be inevitably imprecise, as some members of stereotyped groups may experience more threat than other members of the same groups.¹⁶⁰

Nonetheless, post-hoc adjustments may be attractive to institutions in situations where they cannot adequately reduce stereotype threat. University admissions provides one example. Admissions officers may find that the instruments that best predict subsequent achievement—such as SAT scores and high-school grades—are infected with stereotype threat. It may not be clear how to eliminate stereotype threat from these prior

159. See Valerie Jones Taylor & Gregory M. Walton, *Stereotype Threat Undermines Academic Learning*, 37 PERSONALITY & SOC. PSYCHOL. BULL. 1055, 1064 (2011) (finding that stereotype threat can undermine learning itself, not just performance).

160. Geoffrey L. Cohen & Julio Garcia, “*I Am Us*”: Negative Stereotypes as Collective Threats, 89 J. PERSONALITY & SOC. PSYCHOL. 566, 579 (2005); Rodolfo Mendoza-Denton et al., *Sensitivity to Status-Based Rejection: Implications for African American Students’ College Experience*, 83 J. PERSONALITY & SOC. PSYCHOL. 896, 915–16 (2002); Elizabeth C. Pinel, *Stigma Consciousness: The Psychological Legacy of Social Stereotypes*, 76 J. PERSONALITY & SOC. PSYCHOL. 114, 126 (1999).

contexts, and even if it is clear, the university may be unable to convince testing organizations or high schools to act. In response, officers might decide to account for stereotype threat in evaluating applicants. They may wish to do so in any number of ways, for instance by taking stereotype threat into account as part of an individualized, all-things-considered admissions process, by deemphasizing threat-infected measures like grades and the SAT, or by recalibrating the scores of affected students.

In choosing among these options, a university would be making decisions about cost and predictive power. For instance, individualized admissions processes are usually more expensive than mechanical procedures. And deemphasizing infected measures—for example, by not using SAT scores to choose among those who exceed some minimum score¹⁶¹—could involve a sacrifice in predictive power.¹⁶²

A university that chose to account for stereotype threat through individualized selection or by mechanical recalibration would then face difficult questions associated with implementation. Individuals vary in their experience of stereotype threat, both because of individual differences in susceptibility to the same threat-inducing cues (two students in the same standardized testing environment may experience different levels of threat) and because of differences in the cues people are exposed to (students may attend different high schools that vary in their level of threat). If threat reduces black students' scores on a test by an average of twenty-five points, it will reduce the scores of some black students by more than twenty-five points and others by less. Hence, a correction of twenty-five points for all black students would overcorrect for the effects of stereotype threat in some cases and undercorrect in others. The greater the variation among black students in their experience of stereotype threat, the larger this issue.¹⁶³ This approach would also, of course, involve racial classifications.

Short of eliminating stereotype threat, the best alternative would be to measure each individual student's experience of stereotype threat. Doing so

161. Depending upon the cutoff, the approach would either sweep in unqualified nonstereotyped students or cut out qualified stereotyped students. See, e.g., Walton, Spencer & Erman, *supra* note 14, at 23–24 (noting that the approach would “ignore the continuous nature of applicants’ test scores”). For advocacy of a geography-bound version of this cut-off strategy, see Danielle Allen, *Talent Is Everywhere: Using ZIP Codes and Merit to Enhance Diversity* (with appendix by Tina Eliassi-Rad & Branden Fitelson), in BEYOND AFFIRMATIVE ACTION 145, 145–59 (R. Kahlenberg ed. 2014).

162. See *supra* note 129 and accompanying text (noting the likely lack of less-threatening, similarly predictive measures, and classifying the decision to forego a threatening measure as threat neutralization); notes 93–95, 158–159 (asserting that where feasible, threat neutralization is preferable to corrective measures like those here under consideration).

163. See Feingold, *supra* note 14, at 260 (noting the over- and under-inclusion problem).

would involve no racial classifications¹⁶⁴ and permit more accurate score adjustments than group-level recalibration. An important scientific and policy direction is to develop measures that are adequately valid and reliable to serve this function.¹⁶⁵ For instance, existing measures, which were developed for research purposes, assess individual differences in *susceptibility* to stereotype threat. But because threat emerges from both individual differences (some people experience more threat than others) and *situations* (some situations evoke threat more than others), a valid individual-difference measure of the impact of stereotype threat should reflect both sources of variance.

Further, the fact that existing measures are largely self-reports (e.g., “In school, I worry that people will draw conclusions about my racial group based on my performances”¹⁶⁶) raises four additional complications. First, this approach assumes that people have ready access to their experience of stereotype threat; this may not always be the case, as many psychological phenomena are in part implicit or nonconscious. Second, people may underreport threat, for admitting threat can itself be psychologically threatening.¹⁶⁷ Third, if a self-report measure came to affect admissions decisions, people might be motivated to over-report threat to gain an advantage. Measures used for evaluation and selection must be valid for these purposes and thus reasonably robust to such problems.

Finally, a measure would need to be valid across situations. A common problem with self-report measures is that people answer them in relation to an implicit reference group (e.g., other students at my school). The consequence is that a given measure can be valid within a setting but not across settings. In a recent example, students at several rigorous, high-performing charter schools attended school longer, studied thirty to fifty minutes longer each night, and took on more challenging material than students at a district school.¹⁶⁸ But a battery of self-report assessments showed *no differences* between schools in students’ self-reports of school engagement, self-control, effort, or persistence. Presumably, students in each school answered the self-report questions in reference to their unique

164. *But cf.* Rich, *supra* note 9, at 1560 (arguing that the Supreme Court has repeatedly inferred racial classifications “based on the form and practical effect of facially neutral legislation”).

165. See sources cited *supra* note 160.

166. See Cohen & Garcia, *supra* note 160, at 568.

167. See Logel et al., *supra* note 121, at 300 (noting that activation of a negative stereotype can contribute to subsequent poor performance).

168. See CHRISTINA CLARK TUTTLE ET AL., MATHEMATICA POLICY RESEARCH, KIPP MIDDLE SCHOOLS: IMPACTS ON ACHIEVEMENT AND OTHER OUTCOMES 48 (2013), available at http://www.mathematica-mpr.com/publications/PDFs/education/KIPP_middle.pdf.

comparison group.¹⁶⁹

Addressing these challenges to develop a valid, reliable individual-difference measure of stereotype threat is an important priority. In the meantime, mechanical group-level score recalibration may provide the most accurate way to use threat-infected instruments to measure merit. Where the stereotype driving the threat is a racial one, the recalibration would be race based.

Once a school chooses a means by which it will seek to correct for the effects of stereotype threat, it must decide how extensive a correction is appropriate. Consider a university that has reduced stereotype threat and has no plans to make further reductions. Assume that the university has found that black applicants tend to perform as well, if accepted, as white applicants who score twenty-five points higher on the SAT. Yet suppose research suggests that stereotype threat causes the SAT to underestimate black applicants' potential by an average of fifty points in general. The performance-maximizing university would pursue admissions policies that treat black applicants similarly to white ones who scored twenty-five points higher. The potential-maximizing university would seek to close a fifty-point gap.

B. STRICT SCRUTINY

Correcting for the effects of racial-stereotype threat through mechanical score recalibration or as part of more holistic individualized selection procedures would involve racial classifications subject to strict scrutiny. Because the school would be seeking to make open and competitive measures more accurate, its policy would be permissible only if that end were a compelling governmental interest. But is it? By raising that open question, Part III.B.1 argues, classificatory recalibration of scores illuminates how the Court has structured its equal protection jurisprudence around merit without clearly defining the term or its place in the doctrine.

If the end the school seeks is compelling, Part III.B.2 contends, the resultant inquiry into narrow tailoring would raise additional core equal protection ambiguities. Because narrow tailoring requires that stated purposes be actual purposes, schools must pursue merit as the Court understands it to survive review. But the Court associates merit with both open competition and predictive validity, concerns that may point in opposite directions during admissions processes. According to the Justices,

169. See YEAGER ET AL., *supra* note 101, at 25–26.

narrow tailoring also safeguards equal protection values. And while the Court maintains that those values include preventing racial politics, essentialism, and depreciation and promoting dignity and autonomy, it presumes that racial classifications threaten them all without fully specifying the reach of any. Whether it would extend that presumption to score corrections that use racial classifications to recognize, highlight, and credit individuals' otherwise latent ability remains unclear.

Part III.B concludes with a tentative prediction: The Court is likely to permit score corrections through individualized selection but not through mechanical adjustments.

1. A Compelling Interest in Merit?

To survive application of strict scrutiny, score corrections must serve a compelling governmental interest. The interest must be an as-yet judicially undeclared one, for score corrections do not aim at recognized justifications for racial classifications: remedying an institution's own identifiable past discrimination and achieving broad-based campus diversity. Rather, score corrections aim to improve the accuracy of measures of candidates' underlying potential or likely subsequent performance. Whether such an end qualifies as compelling will likely turn on whether the Court recognizes one or the other to be meritocratic.¹⁷⁰ That is so, this section argues, because the Court is likely—though not certain—to find a compelling interest in merit, as the Justices define it.

Recognition of meritocracy as compelling would reflect its place of privilege in U.S. life. It is a too-often-unfulfilled truism that people should be judged on their merits—which are presumed to be individual attributes—rather than on arbitrary traits.¹⁷¹ We look with reverence upon

170. Abraham L. Wickelgren, *Affirmative Action: More Efficient than Color Blindness*, 10 TEX. J. C.L. & C.R. 165 (2005) (exploring conditions under which a self-interested entity would seek to select applicants of color with slightly lower scores on prior measures); Yoon-Ho A. Lee, Social Perception and Affirmative Action: A Game-Theoretic Analysis (Dec. 19, 2007) (unpublished manuscript), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1903601 (similar).

171. Michael H. Shapiro, *The Impact of Genetic Enhancement on Equality*, 34 WAKE FOREST L. REV. 561, 586–87 (1999). See also, e.g., JOHN CARSON, THE MEASURE OF MERIT: TALENTS, INTELLIGENCE, AND INEQUALITY IN THE FRENCH AND AMERICAN REPUBLICS, 1750–1940 *passim* (2007) (tracing conceptions of merit as an individual attribute from the mid-eighteenth to the mid-twentieth century, and describing how standardized tests were used to justify racial and other exclusions); ROGERS M. SMITH, CIVIC IDEALS: CONFLICTING VISIONS OF CITIZENSHIP IN U.S. HISTORY (1999) (describing ascriptive strands in U.S. civic identity); Michael H. Shapiro, *Who Merits Merit? Problems in Distributive Justice and Utility Posed by the New Biology*, 48 S. CAL. L. REV. 318, 321–24 (1974) (describing how biological enhancements are likely to raise difficult questions for those seeking to treat merit as individual attributes).

the dream that people “not be judged by the color of their skin, but by the content of their character.”¹⁷² Such aspirations characterized the postwar turn in college admissions toward standardized testing, which advocates cast as a shift to selection based on ability rather than privilege.¹⁷³ Along similar lines, the Court has noted that even legal preferences—like those for veterans—“represent an awkward—and, many argue, unfair—exception to the widely shared view that merit and merit alone should prevail in the employment policies of government.”¹⁷⁴ One reason that many people perceive racial discrimination to be odious is that it generally violates the merit principle.¹⁷⁵ Meritocratic decision making, conversely, is often seen as consistent with—even constitutive of—equality and fairness.¹⁷⁶ President Bill Clinton expressed this sentiment when he argued that to achieve “fairness,” “affirmative action has to be made consistent with our highest ideals of . . . merit.”¹⁷⁷

Merit stalks equal protection jurisprudence. It is a shadow interest, treated as compelling but as yet undeclared. That ambiguous place makes it of a piece with the Court’s resolutely unstructured compelling-interest jurisprudence. Yet, it is also distinguishable, the rare potential interest for which the Court has laid groundwork for extending express recognition.

Having “frequently adopted” a “casual approach,”¹⁷⁸ the Court has left few clues for those seeking to determine whether particular governmental

172. *Excerpts from Addresses at Lincoln Memorial During Capital Civil Rights March*, N.Y. TIMES, Aug. 29, 1963, at 21 (quoting Martin Luther King, Jr.).

173. See, e.g., Michael Ackerman, *Mental Testing and the Expansion of Educational Opportunity*, 35 HIST. EDUC. Q. 279, 280 (1995) (“[A]dvocates of equal educational opportunity argued that a meritocratic testing program could help reverse the harmful consequences of . . . discrimination. In addition, educators who favored the expansion of higher education maintained that enrollments could be increased by drawing upon previously untapped talent with the aid of newly developed measures of ability and aptitude.”). For a fuller discussion of the postwar embrace of standardized testing in college admissions, see generally NICHOLAS LEMANN, *THE BIG TEST: THE SECRET HISTORY OF THE AMERICAN MERITOCRACY* (1999).

174. *Pers. Adm'r of Mass. v. Feeney*, 442 U.S. 256, 280 (1979).

175. E.g., Michael H. Shapiro, *Does Technological Enhancement of Human Traits Threaten Human Equality and Democracy?*, 39 SAN DIEGO L. REV. 769, 791–92 (2002); D. Ramona Bobocel et al., *Justice-Based Opposition to Social Policies: Is It Genuine?*, 75 J. PERSONALITY & SOC. PSYCHOL. 653, 654, 667 (1998).

176. See, e.g., Michael H. Shapiro, *The Identity of Identity: Moral and Legal Aspects of Technological Self-Transformation*, 22 SOC. PHIL. & POL'Y 308, 308 (2005) (“For many, keying the size of rewards to degree of merit . . . does not violate standards of equality or fairness, and may indeed promote them.”).

177. William J. Clinton, Remarks on Affirmative Action at the National Archives and Records Administration (July 19, 1995), in 2 PUBLIC PAPERS OF THE PRESIDENTS OF THE UNITED STATES: WILLIAM J. CLINTON 1106, 1113 (1997).

178. Fallon, *supra* note 45, at 1321.

interests count as compelling beyond acknowledging that some may remain to be declared.¹⁷⁹ Scholars have identified little coherent pattern in the interests that the Court has recognized.¹⁸⁰ It is hard to know what unites the two recognized interests in the race-discrimination context, which are remedying identifiable past discrimination and pursuing diversity in higher education. This is especially so given that the seemingly weightier and more constitutionally relevant one of remedying discrimination has been all but excised in favor of the upstart alternative of attaining diversity.¹⁸¹

Although confident predictions are hazardous, the Court has expressed solicitude for merit consonant with its pursuit being a compelling governmental interest, repeatedly stating, for instance, that it “demeans the dignity and worth of a person to be judged by ancestry instead of by his or her own merit.”¹⁸² Similarly, the decisive opinions in *Regents of the*

179. *Id.* at 1321–24. See also *Grutter v. Bollinger*, 539 U.S. 306, 328 (2003) (stating, in the course of recognizing academic diversity as a compelling interest, that “we have never held that the only governmental use of race that can survive strict scrutiny is remedying past discrimination.”); Stephen E. Gottlieb, *Tears for Tiers on the Rehnquist Court*, 4 U. PA. J. CONST. L. 350, 366 (2002) (arguing that the Rehnquist Court “has never moved beyond an ad hoc specification of the elements of the compelling interest test”); Stephen E. Gottlieb, *Compelling Governmental Interests: An Essential but Unanalyzed Term in Constitutional Adjudication*, 68 B.U. L. REV. 917, 937 (1988) [hereinafter Gottlieb, *Compelling Interests*] (“[W]ith few exceptions, the Court has failed to explain the basis for finding . . . compelling governmental interests.”).

180. E.g., Fallon, *supra* note 45, at 1321–25; Gottlieb, *Compelling Interests*, *supra* note 179, at 932–37.

181. See *Fisher v. Univ. of Tex. at Austin*, 133 S. Ct. 2411, 2423 (2013) (“[T]he Court has recognized that the government has a compelling interest in remedying past discrimination for which it is responsible, but we have stressed that a government wishing to use race must provide a strong basis in evidence for its conclusion that remedial action [is] necessary.” (alteration in original) (internal quotation marks omitted)); *Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 720–21 (2007) (similar).

182. *Rice v. Cayetano*, 528 U.S. 495, 517 (2000). See also *Gratz v. Bollinger*, 539 U.S. 244, 271 (2003) (“Justice Powell’s opinion in *Bakke* emphasized the importance of considering each particular applicant as an individual, assessing all of the qualities that individual possesses, and in turn, evaluating that individual’s ability to contribute to the unique setting of higher education.”); *Shaw v. Reno*, 509 U.S. 630, 643 (1993) (“[A]n explicit policy of assignment by race may . . . suggest[] the utility and propriety of basing decisions on a factor that ideally bears no relationship to an individual’s worth” (quoting *United Jewish Orgs. of Williamsburgh, Inc. v. Carey*, 430 U.S. 144, 173 (1977) (Brennan, J., concurring in part))); *DeFunis v. Odegaard*, 416 U.S. 312, 337 (1974) (Douglas, J., dissenting) (“A DeFunis who is white is entitled to no advantage by reason of that fact; nor is he subject to any disability, no matter what his race or color. Whatever his race, he had a constitutional right to have his application considered on its individual merits in a racially neutral manner.”). Cf. *Grutter*, 539 U.S. at 339–40 (“Narrow tailoring does not require . . . a university to choose between maintaining a reputation for excellence or fulfilling a commitment to provide educational opportunities to members of all racial groups. . . . [It requires no] dramatic sacrifice of diversity, the academic quality of all admitted students, or both. . . . [, and permits the university to decline] to abandon the academic selectivity that is the cornerstone of its educational mission.”); Siegel, *supra* note 139, at 810 (“[W]hen consideration of individual merit is possible and the Court regards the social benefits of using race as small, the merit

University of California v. Bakke, *Gratz v. Bollinger*, *Grutter v. Bollinger*, and *Fisher v. University of Texas*¹⁸³ all criticized racial classifications that “insulat[e] [candidates] . . . from competition” with one another.¹⁸⁴ So too did *Ricci v. DeStefano*¹⁸⁵ when, speaking of Title VII in a constitutional register, it condemned abandoning what it cast as apparently meritocratic test results.¹⁸⁶ Absent resort to race, the Court appears to approvingly presume and at times explicitly states that decision makers “naturally would focus on the qualifications.”¹⁸⁷ In rare cases in which the Court recognizes that the absence of racial classifications would undermine merit, it permits limited racial classifications. Thus, in *Grutter*, the Court allowed pursuit of diversity in higher education through the use of limited racial classifications notwithstanding the availability of admissions by lottery as an alternative.¹⁸⁸

factor can be important.”).

183. Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265 (1978); *Gratz v. Bollinger*, 539 U.S. 244 (2003), *Grutter v. Bollinger*, 539 U.S. 306 (2003); *Fisher v. Univ. of Texas at Austin* 133 S. Ct. 2411 (2013).

184. *Grutter*, 539 U.S. at 334 (quoting *Bakke*, 438 U.S. at 315 (Powell, J.)); *Gratz*, 539 U.S. at 271–74; *Fisher*, 133 S. Ct. at 2418.

185. *Ricci v. DeStefano*, 557 U.S. 557 (2009).

186. See *infra* notes 217–224, 230–235 and accompanying text. Scholars have noted—and questioned—the portrayal. See Mark S. Brodin, *Ricci v. DeStefano: The New Haven Firefighters Case & the Triumph of White Privilege*, 20 S. CAL. REV. L. & SOC. JUST. 161, 163 (2011) (“Ricci . . . redefined merit by equating it with success on multiple-choice examinations, as opposed to more reliable methods of personnel selection.”); *id.* at 202–23, 229; Cheryl I. Harris & Kimberly West-Faulcon, *Reading Ricci: Whitenig Discrimination, Racing Test Fairness*, 58 UCLA L. REV. 73, 84 (2010) (“[T]he decision ignores the central question[:] . . . Did candidates’ scores on the tests correlate to job performance and thus identify the best candidates for the job?”); Helen Norton, *The Supreme Court’s Post-Racial Turn Towards a Zero-Sum Understanding of Equality*, 52 WM. & MARY L. REV. 197, 226 (2010) (“Considerable uncertainty remained, however, over the tests’ accuracy in identifying the best candidates for promotion and the possibility of less discriminatory alternatives that better predict successful performance in leadership positions. In invalidating New Haven’s response to such uncertainty, the Court simply denied its existence.”); Cedric Merlin Powell, *Harvesting New Conceptions of Equality: Opportunity, Results, and Neutrality*, 31 ST. LOUIS U. PUB. L. REV. 255, 318 (2012) (“Ricci’s failure to apply Title VII’s requirements regarding test validation actually enacts a presumption that white overrepresentation is the natural product of merit selection” (quoting Harris & West-Faulcon, *supra*, at 157)); Rich, *supra* note 34, at 77.

187. *Price Waterhouse v. Hopkins*, 490 U.S. 228, 243 (1989) (plurality opinion), quoted in Rich, *supra* note 34, at 57. See also Siegel, *supra* note 139, at 810 (noting that a key factor in affirmative action cases is whether something like merit selection—for instance, a lowest bid in a contracting case or grades and test scores in university admissions—is possible); Kimberly Jenkins Robinson, *The Constitutional Future of Race-Neutral Efforts to Achieve Diversity and Avoid Racial Isolation in Elementary and Secondary Schools*, 50 B.C. L. REV. 277, 347 (2009) (“The Supreme Court has explained that racial classifications are forbidden because judging someone by his or her race undermines the worth and dignity of individuals when personal qualities and merit represent the appropriate measure.”).

188. *Grutter*, 539 U.S. at 334, 340. See also *id.* at 392–93 (Kennedy, J., dissenting) (“There is no constitutional objection to the goal of considering race as one modest factor among many others to

The Court is especially likely to recognize merit in the educational context.¹⁸⁹ As *Fisher* stated, because the “academic mission of a university is a special concern of the First Amendment,” a university’s “educational judgment that . . . diversity is essential to its educational mission”—as the Court appeared to accept as true for Michigan Law School—“is one to which we defer.”¹⁹⁰ On review, courts “ensure that there is a reasoned, principled explanation for the academic decision.”¹⁹¹ Were a similarly deferential standard to apply to a university’s—or public school’s—judgment that merit was essential to its educational mission, merit would likely qualify as a compelling governmental interest.

As with diversity, merit is not self-defining.¹⁹² In the former area, the Court has recognized only broad-based diversity—to which racial heterogeneity makes but a contribution—as a compelling interest.¹⁹³ The remainder of this section seeks to draw from the Court’s informal treatment of merit clues as to which visions of the concept the Court is most likely to embrace. It argues that the Court understands meritocratic decision making to result from open, competitive processes that seek to sort candidates based on their predicted subsequent performance if selected. Subsequent performance can, of course, be measured in many ways—future grades, contributions to classroom dynamics, leadership and civic-mindedness following graduation, etc.—not all of which the Court (or every school) values equally. In exploring these issues, it is helpful to distinguish between what selecting organizations value—for example, likely subsequent performance, defined in any number of ways¹⁹⁴—and the

achieve diversity . . .”); *Fisher*, 133 S. Ct. at 2419–20 (similar); *Parents Involved*, 551 U.S. at 790 (Kennedy, J., concurring in part and concurring in the judgment) (similar, but for public schools); Ozan O. Varol, *Strict in Theory, but Accommodating in Fact?*, 75 MO. L. REV. 1243, 1261–62 (2010) (The *Grutter* Court “deferred . . . to the law school’s interest in maintaining its elite status.”).

189. For the possibility that the interest in diversity extends beyond higher education to all educational institutions, see *Parents Involved*, 551 U.S. at 797 (Kennedy, J., concurring in part and concurring in the judgment) (“A compelling interest exists in avoiding racial isolation, an interest that a school district, in its discretion and expertise, may choose to pursue.”); *id.* at 790 (indicating that public schools may be able to employ racial classifications under certain circumstances); and *id.* at 788–89 (finding it permissible for public schools to use race-conscious measures to seek a diverse student body).

190. *Fisher*, 133 S. Ct. at 2418, 2419 (internal quotation marks and citations omitted). See also Pamela S. Karlan, *Compelling Interests/Compelling Institutions: Law Schools as Constitutional Litigants*, 54 UCLA L. REV. 1613, 1621–22 (2007) (noting that the Court “declares that racial diversity is compelling because a school thinks it is”).

191. *Fisher*, 133 S. Ct. at 2419.

192. For an overview of additional concerns underlying notions of merit, see, for example, GEORGE SHER, *DESERT* (1987).

193. E.g., *Grutter*, 539 U.S. at 327–42.

194. One might also value likely subsequent potential: one’s total ability—latent and expressed—

instruments that they rely upon to do so—such as prior test scores and grades.

As to what the Court envisions meritocratic organizations measuring, the employment context provides a helpful starting point. There, the Court lauds antidiscrimination law for permitting employment decisions to be based on the value that an employee is likely to create for her employer.¹⁹⁵ Thus, the Court has reasoned, veteran preferences are non-meritocratic because they do “not purport to define a job-related characteristic,”¹⁹⁶ while Title VII’s anti-discriminatory provisions aim “to promote hiring on the basis of job qualifications.”¹⁹⁷ In the case of firefighter promotions, the Court dramatically quoted a litigant who argued against abandoning apparently non-discriminatory test results: “When your life’s on the line, second best may not be good enough.”¹⁹⁸ And in hiring a contractor, the Court has indicated, basing the choice on expected price rather than race promotes decision making on the merits.¹⁹⁹

In the education context, the Court conceives of subsequent performance in two ways: benefit to the school and achievement of the individual. The Court has cast the importance of diversity to schools in the former terms: It helps them promote “speculation, experiment and creation”,²⁰⁰ a “robust exchange of ideas”,²⁰¹ “cross-racial understanding” and the “break down [of] racial stereotypes”; “livelier, more spirited, and . . . more enlightening and interesting” “classroom

whatever the probability that the latent ability would eventually be expressed. As will become evident, that has not been the Court’s approach.

195. See, e.g., Ricci v. DeStefano, 557 U.S. 557, 628 (2009) (“The very purpose of [Title VII] is to ensure that individuals are hired and promoted based on qualifications manifestly necessary to successful performance of the job in question, qualifications that do not screen out members of any race.”); Robert C. Post, *Prefjudicial Appearances: The Logic of American Antidiscrimination Law*, 88 CAL. L. REV. 1, 11, 15 (2000) (“Blindness” to race sometimes required by antidiscrimination laws “requires employers to base their judgments instead upon . . . ‘individual merit’ or ‘intrinsic worth,’” a theory that presumes “that successful job performance is . . . distinct from . . . race.”); David B. Cruz, *Making Up Women: Casinos, Cosmetics, and Title VII*, 5 NEV. L.J. 240, 251–55 (2004) (critiquing Post’s critique); Jeb Rubenfeld, *The Right of Privacy and the Right to Be Treated as an Object*, 89 GEO. L.J. 2099, 2100 (2001) (portraying Title VII as causing employers to treat employees and job candidates “as embodied net marginal product”).

196. Pers. Adm’r of Mass. v. Feeney, 442 U.S. 256, 277 (1979); *id.* at 280.

197. *Ricci*, 557 U.S. at 582 (quoting *Griggs v. Duke Power Co.*, 401 U.S. 424, 434 (1971)).

198. *Id.* at 568 (quoting Petition for Certiorari app. at A787–A788, *Ricci v. DeStefano*, 557 U.S. 557 (2009) (No. 06–4996–CV) [hereinafter *Ricci* Petition Appendix]).

199. *Adarand Constructors, Inc. v. Pena*, 515 U.S. 200, 205, *passim* (1995).

200. *Fisher v. Univ. of Tex. at Austin*, 133 S. Ct. 2411, 2418 (2013) (quoting *Sweezy v. New Hampshire*, 354 U.S. 234, 263 (1957) (Frankfurter, J., concurring in the judgment)).

201. *Grutter v. Bollinger*, 539 U.S. 306, 324 (2003) (quoting *Regents of the Univ. of Cal. v. Bakke*, 438 U.S. 265, 312 (1978) (Powell, J.))

discussion”; and better preparation of students “for an increasingly diverse workforce and society.”²⁰² While meritocratic score corrections would also likely result in greater diversity and in stronger students who would improve classroom discussion, that is not the end that they are designed to achieve.²⁰³

The Court also—and particularly—values what some score corrections do pursue: selection of students most likely to perform well on academic measures at the admitting school.²⁰⁴ (By contrast, the Court has displayed little solicitude for potential unlikely to be realized.) When approving affirmative-action programs, for instance, the Court stresses that such programs only benefit qualified and “admissible” students “capable of doing good work in their courses.”²⁰⁵ Similarly, schools may use racial classifications to pursue diversity even when they could achieve similar ends by abandoning “academic selectivity” and “a reputation for excellence” founded upon the “academic quality of all admitted students.”²⁰⁶ One reason that the Court scrutinizes affirmative action is because it selects on grounds other than predicted academic achievement. As Justice Lewis Powell explained in his *Bakke* concurrence, which is now a touchstone of the Court’s diversity jurisprudence²⁰⁷: “[C]onnected with

202. *Id.* at 330 (internal quotation marks omitted). Compare *id.* at 331–33 (portraying the university as interested in ensuring that the path to post-graduation leadership remains open), *with Fisher*, 133 S. Ct. at 2417–18 (limiting school’s interest in diversity to campus dynamics), *and with* WILLIAM G. BOWEN & DEREK BOK, THE SHAPE OF THE RIVER: LONG-TERM CONSEQUENCES OF CONSIDERING RACE IN COLLEGE AND UNIVERSITY ADMISSIONS (1998) (valuing post-graduation leadership, community service, and civic behavior above grades).

203. *Grutter*, 539 U.S. at 338 (“By virtue of our Nation’s struggle with racial inequality, such [“underrepresented minority”] students are both likely to have experiences of particular importance to the Law School’s mission, and less likely to be admitted in meaningful numbers on criteria that ignore those experiences.”). One might aim to make a form of score corrections part of an attempt to achieve diversity and thereby further twine that compelling interest together with merit. *See supra* note 188 and accompanying text. This Article does not take up the distinct and interesting question of whether such an approach to diversity would be constitutionally preferred.

204. That schools would likely implement score corrections for this reason does not mean that they would seek to select students based wholly or even primarily on their likely future performance on academic measures. Rather, it means that the point of the score correction is to permit the school to make a more accurate judgment about what those grades are likely to be—however they are then weighed in the admissions process.

205. *Bakke*, 438 U.S. at 316 (Powell, J.) (quoting Brief for Columbia University, Harvard University, Stanford University and the University of Pennsylvania as Amici Curiae app. at 2, Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265 (1978) (No. 76–811)), quoted in *Grutter*, 539 U.S. at 338.

206. *Grutter*, 539 U.S. at 339–40 (describing concern with academic performance and standards as “the cornerstone of [the] educational mission” of Michigan Law School).

207. *See id.* at 323 (“Since this Court’s splintered decision in *Bakke*, Justice Powell’s opinion announcing the judgment of the Court has served as the touchstone for constitutional analysis of race-conscious admissions policies.”); *id.* at 322, 324–25, 329–30, 334–37, 339, 341, 343 (relying on Powell’s opinion in *Bakke*); *Fisher*, 133 S. Ct. at 2415, 2417–22 (same).

the idea of preference itself” is “inequity in forcing innocent persons . . . to bear the burdens of redressing grievances.”²⁰⁸ By contrast, he wrote, “[r]acial classifications” that “serve [the] purpose [of] fair appraisal of . . . academic promise . . . might” involve “no ‘preference’ at all” and, one presumes, no concomitant inequity.²⁰⁹

In addition to expressing views on the content of meritocracy, the Court has also suggested the process through which it should be conducted: an open and competitive process.²¹⁰ In the context of public contracting, for example, it has struck down racially classificatory affirmative-action plans that wholly deny “certain citizens the opportunity to compete for” certain projects.²¹¹ Affirmative-action policies that use racial classifications to promote diversity, the Justices have also insisted, may not “insulat[e] each category of applicants with certain desired qualifications from competition with all other applicants.”²¹² Instead, they must place all applicants “on the same footing for consideration”²¹³ and ensure that “each candidate ‘compete[s] with all other qualified applicants.’”²¹⁴ Justice Powell elaborated the point in *Bakke*. Programs that do not ensure competition “will be viewed as inherently unfair by the public generally” because “[f]airness in individual competition for opportunities . . . is a widely cherished American ethic.”²¹⁵ But under programs that do permit all comers to compete with each other, he added, each applicant’s “qualifications would have been weighed fairly and competitively,” which would leave “no basis to complain of unequal treatment under” equal protection.²¹⁶

208. *Bakke*, 438 U.S. at 298 (Powell, J.); Siegel, *supra* note 9, at 39.

209. *Bakke*, 438 U.S. at 306 n.43 (Powell, J.). See also *id.* at 296–97 nn.36–37; Siegel, *supra* note 139, at 809–11. Cf. *United Steelworkers of Am. v. Weber*, 443 U.S. 193 (1979) (addressing similar issues under Title VII); *Johnson v. Transp. Agency, Santa Clara Cnty., Cal.* 480 U.S. 616 (1987) (same).

210. See Richard M. Re, *Relative Standing*, 102 GEO. L.J. 1191, 1201 (2014) (observing how the Court stretches the standing doctrine to count having been “prevent[ed] from competing on an equal footing” as injury in fact (alteration in original) (quoting *Adarand Constructors, Inc. v. Pena*, 515 U.S. 200, 211 (1995))).

211. *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 493 (1989) (plurality opinion).

212. *Grutter*, 539 U.S. at 334 (alteration in original) (quoting *Bakke*, 438 U.S. at 315 (Powell, J.)). See also *Bakke*, 438 U.S. at 317 (adding that permissible plans may not “insulate the individual from comparison with all other candidates for the available seats”).

213. *Grutter*, 539 U.S. at 334 (quoting *Bakke*, 438 U.S. at 317 (Powell, J.)).

214. *Grutter*, 539 U.S. at 335 (quoting *Johnson*, 480 U.S. at 638); *id.* at 337.

215. *Bakke*, 438 U.S. at 319 n.53 (Powell, J.).

216. *Id.* at 318. Cf. *Grutter*, 539 U.S. at 389 (Kennedy, J., dissenting) (criticizing an admissions program that he cast as making “race . . . likely outcome determinative for many members of minority groups . . . where the competition becomes tight”—in seeking one of the “15% to 20% of the seats” in a class that go to students without top test scores and grades). The Court depicts competitions as

The Court is often protective of standardized tests and grades, which it views as generally open, competitive, and relatively accurate predictors of subsequent performance.²¹⁷ In *Ricci v. DeStefano*, for instance, the Justices deemed it to be illegal discrimination when New Haven, after learning that white firefighters had scored disproportionately well on a test to determine eligibility for promotions, threw out the results.²¹⁸ The Court declared that New Haven law “establishe[d] a merit system” that, in promotion cases, relied “on objective examinations to identify the best qualified candidates.”²¹⁹ Having sat for the competitive test, the Court explained, candidates had “legitimate expectations” to have the results honored.²²⁰ And while the test was open to firefighters eligible for the promotions, success could require “considerable personal and financial expense” that—once overcome—rendered any retesting requirement “all the more” objectionable.²²¹

The Court cast *Ricci* as “also . . . in keeping with Title VII’s” permission to employers under § 2000e-2(h) to “act upon the results of any professionally developed ability test provided that . . . its . . . action upon the results is not . . . used to discriminate.”²²² Prior cases had emphasized that § 2000e-2(h) “did not alter the meaning of Title VII,” which elsewhere barred unjustified employment practices with disparate racial impacts.²²³ But now the Court perceived in the language of § 2000e-2(h) “express protection of bona fide promotional examinations,”²²⁴ the

involving investments of effort, and casts the failure to honor that effort as part of the harm results from unfairly withholding promised prizes from rightful winners. *See Ricci v. DeStefano*, 557 U.S. 557, 562, 564 (2009) (invalidating city’s decision to throw out results of a “merit system” promotional exam for which “[m]any firefighters studied for months, at considerable personal and financial cost”); *Croson*, 488 U.S. at 481–83, *passim* (describing a contractor’s efforts to create a winning bid for a city contract for which no other bids were received, and striking down the affirmative action law that would have otherwise denied the bidder the contract). Cf. SHER, *supra* note 192, at 53–68 (discussing diligence as a basis for merit). The Court treats test scores and grades as satisfying this criterion.

217. Cf. Michelle Richardson, Charles Abraham & Rod Bond, *Psychological Correlates of University Students’ Academic Performance: A Systematic Review and Meta-Analysis*, 138 PSYCHOL. BULL. 353, 354 (2012) (“[H]igh school GPA is a stronger predictor of university GPA than is either the SAT or the ACT. All three measures have been found to explain independent variation in GPA . . . ”).

218. *Ricci*, 557 U.S. at 562, 583–84 (reasoning that once employers administer examinations to decide the “intense competition for promotions,” they “create legitimate expectations on the part of those who took the tests” and thus sometimes must honor those results even if they produce racially disparate outcomes).

219. *Id.* at 564 (adding that it is a “merit system” to “fill vacancies . . . with the most qualified individuals, as determined by job-related examinations”).

220. *Id.* at 583.

221. *Id.* at 593.

222. *Id.* at 584 (quoting 42 U.S.C. § 2000e-2(h) (2012)).

223. *Connecticut v. Teal*, 457 U.S. 440, 452 (1982) (internal quotation marks omitted).

224. *Ricci*, 557 U.S. at 584.

suggestion both that Congress equated use of such examinations with hiring on the basis of job qualifications and that the practice was in fact meritocratic.

Justices have displayed similar faith in tests and grades in the educational context. In *Bakke*, Justice Powell dismissed the contention that “but for pervasive racial discrimination” minority applicants would have qualified for admissions under existing measures at rates similar to those produced by an affirmative-action program.²²⁵ Chastising proponents of the theory because “[n]ot one word in the record supports this conclusion,”²²⁶ he presumed that traditional bases for admission would generally provide a “fair appraisal of each individual’s academic promise.”²²⁷ *Grutter* echoed Powell’s assessment, largely equating academic selectivity—primarily based on grades and LSAT scores—with “academic quality.”²²⁸ Like *Gratz* and *Fisher*, it also portrayed admissions processes that depend heavily on grades and test scores as open and competitive in ways that the introduction of racial classifications might undermine.²²⁹

Yet, the Justices have remained sensitive to the possibility that selecting organizations may seek to improve tests to level competitive playing fields and increase their predictive power. As already mentioned, the Court along with most past social scientific research has tended to focus on the content of tests more so than their context. *Ricci*, for instance, reassured employers that they could “design [a] test . . . to provide a fair opportunity for all individuals, regardless of their race,” by “invit[ing] comments” and undertaking “open discussions toward that end.”²³⁰ New Haven had taken some such steps, the Court observed, before perhaps too hastily declaring them successful. The city had “thought about promotion qualifications and relevant experience in [race] neutral ways,” the Court recounted, and been “careful to ensure broad racial participation in the design of the test itself and its administration.”²³¹ Test questions were chosen and reviewed by methods likely to guard against discriminatory

225. Regents of the Univ. of Calif. v. Bakke, 438 U.S. 265, 296 n.36 (1978) (Powell, J.) (quoting *Bakke*, 438 U.S. at 365–66 (Brennan, White, Marshall & Blackmun, JJ., concurring in the judgment in part and dissenting in part)).

226. *Id.*

227. *Id.* at 306 n.43.

228. *Grutter v. Bollinger*, 539 U.S. 306, 340 (2003).

229. *Id.* at 334–41; *Gratz v. Bollinger*, 539 U.S. 244 (2003); *Fisher v. Univ. of Tex. at Austin*, 133 S. Ct. 2411 (2013).

230. *Ricci v. DeStefano*, 557 U.S. 557, 585 (2009).

231. *Id.* at 593.

content, it added.²³² And presumably, for the same reason, it related, interviews were conducted by racially diverse panels of unbiased expert outsiders.²³³ In the face of criticism of the content of the test, the Court insisted that the exam's design and administration were "open and fair" and recounted how a creator of the test had "implor[ed] anyone that had . . . concerns to review the content of the exam."²³⁴ The chairman of the New Haven Civil Service Board had observed that "nobody convinced me that we can feel comfortable that, in fact, there's some likelihood that there's going to be an exam designed that's going to be less discriminatory."²³⁵

In *Bakke*, Justice Powell contemplated the implications of using a measure infected by a racial bias. Apparently assuming that schools might use such biased measures, Justice Powell perceived a choice between recalibrated or biased scores in which the former would likely be permissible:

Racial classifications in admissions conceivably could serve [the] . . . purpose [of] . . . fair appraisal of each individual's academic promise in the light of some cultural bias in grading or testing procedures. To the extent that race and ethnic background were considered only to the extent of curing established inaccuracies in predicting academic performance, it might be argued that there is no "preference" at all.²³⁶

In essence, Justice Powell argued that if students are to be admitted, schools would be justified in choosing those students expected to perform best.

Given its solicitude for processes that it deems meritocratic, the Court is likely, but not certain, to recognize merit (or a related interest, such as ensuring that otherwise meritocratic selection procedures do not produce unjustified racial disparities) as compelling. If so, it will be merit as the Court defines it: selecting for likely subsequent academic performance through an open and competitive process.

232. *Id.* at 564–65 ("At every stage of the job analyses, [the test designer], by deliberate choice, oversampled minority firefighters to ensure that the results—which [the designer] would use to develop the examinations—would not unintentionally favor white candidates."); *id.* at 566–74.

233. *Id.* at 564–66.

234. *Id.* at 593 (internal quotation marks omitted); *id.* at 569.

235. *Id.* at 574 (quoting Ricci Petition Appendix, *supra* note 198, at A1159–A1160).

236. Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265, 306 n.43 (1978) (Powell, J.). See also *id.* at 296–97 nn.36–37; Feingold, *supra* note 14, at 231–32, 234, 256 (advocating for a compelling interest much like merit that would be a hybrid of diversity and remedying discrimination).

2. Narrowly Tailored to What End?

Even if merit serves as a compelling governmental interest, use of racial classifications to correct for the effects of stereotype threat will only survive strict scrutiny when narrowly tailored to this end.²³⁷ While conclusions as to what such tailoring would require can only be speculative, the Court's diversity jurisprudence provides important clues. It suggests that some approach to score corrections is likely permissible, but that determining which is fraught due to ambiguities concerning the goals of narrow tailoring, the contours of equal protection values, and the meaning of merit. Teasing apart these aspects of narrow tailoring yields few clear answers beyond an educated guess that accounting for the effects of stereotype threat through individualized assessment is more likely to pass constitutional and Title VI muster than mechanical score recalibration.

Under existing diversity jurisprudence, key factors determining the permissibility of resort to racial classifications include availability of nonclassificatory alternatives and temporal scope. Universities may only use racial classifications, the Court proclaims, absent reasonably effective nonclassificatory alternatives. In meeting that standard, educational institutions must investigate nonclassificatory alternatives and convince reviewing courts that no adequate substitute to racial classifications exists.²³⁸ To be adequate, the Court explains, a “nonracial approach” need only “promote the substantial interest about as well” as the racial classification for which it would substitute.²³⁹ Where that is so, the substitute must be used notwithstanding any additional “tolerable administrative expense.”²⁴⁰ The primary candidates for adequate

237. *Cf. Fisher v. Univ. of Texas at Austin*, 133 S. Ct. 2411, 2419 (2013) (emphasizing that “[s]trict scrutiny is a searching examination”); Tomiko Brown-Nagin, *Rethinking Proxies for Disadvantage in Higher Education: A First Generation Students’ Project*, 2014 U. CHI. LEGAL F. 433, 460 (2014) (“Fisher . . . made clear that universities must surmount a substantial evidentiary burden before turning to race-conscious admissions policies” in lieu of “race-neutral alternatives . . . ”).

238. *Fisher*, 133 S. Ct. at 2420. *See also Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 790 (2007) (Kennedy, J., concurring in part and concurring in the judgment) (“[I]ndividual racial classifications employed in this manner may be considered legitimate only if they are a last resort to achieve a compelling interest.”); Siegel, *supra* note 139, at 796 (“[R]ace-conscious affirmative action programs . . . may be used only after a ‘serious, good faith consideration of workable race-neutral alternatives . . . ’” (quoting *Grutter v. Bollinger*, 539 U.S. 306, 339 (2003))). For possible non-classificatory alternatives, see Allen, *supra* note 161; Richard D. Kahlenberg, *What Sotomayor Gets Wrong About Affirmative Action*, CHRON. HIGHER EDUC. (June 17, 2014), <http://chronicle.com/article/What-Sotomayor-Gets-Wrong/147169/>.

239. *Fisher*, 133 S. Ct. at 2420 (internal quotation marks omitted).

240. *Id.*; *Gratz v. Bollinger*, 539 U.S. 244, 275 (2003) (“[T]he fact that the implementation of a program capable of providing individualized consideration might present administrative challenges does not render constitutional an otherwise problematic system.”); *Parents Involved*, 551 U.S. at 788–

substitution are (1) eliminating threat from measures, (2) using only measures uninfected by threat,²⁴¹ and (3) restricting the use of threat-infected measures. As already explained, these approaches should be adopted where effective, but in some cases will be unavailable and thus not achieve merit about as well as recourse to racial classifications.²⁴²

Even when otherwise acceptable, the Court instructs, “race-conscious admissions policies must be limited in time” to survive strict scrutiny.²⁴³ This requirement accords with Justice Kennedy’s view that racial inequality should be treated as a solvable—but as yet unsolved—problem. Score recalibration qualifies because it has a definite stopping point. It arises as institutions seek to reduce stereotype threat. It stops when more institutions reduce threat, thereby reducing the need for score recalibration.²⁴⁴

One might object that perhaps many organizations will never aim to reduce stereotype threat as far as they might. Here, it is worth recalling that the circumstances under consideration rest on a prediction about the future. Institutions are only likely to consider score recalibration once they have made sustained and measurable reductions in stereotype threat. Whether they can achieve such reductions is a question now being tested.²⁴⁵ In the interim, many people are skeptical that short, cheap interventions can make a substantial dent in the seemingly intractable problem of racial inequality.²⁴⁶ But as rigorous randomized controlled studies increasingly show that this is possible, many institutions will seek these low-cost benefits and avoid the stigma of indifference to racialized threat. In a world

90, 796–98 (Kennedy, J., concurring in part and concurring in the judgment) (envisioning creative, diligent exhaustion of alternatives prior to resort to racial classifications).

241. Cf., e.g., *Grutter*, 539 U.S. at 361–62 (Thomas, J., dissenting) (advocating that schools that seek diversity take such an approach).

242. See *supra* Parts I.B.2, III.A. The requirement to avoid racial classifications where alternatives are “about as” effective has some bite. To conform, schools may sometimes have to forego racial classifications that they would otherwise use to make modest gains in terms of cost or accuracy.

243. *Grutter*, 539 U.S. at 309; Siegel, *supra* note 139, at 796 & n.66. Although Justice Kennedy did not join the majority in *Grutter*, his views on this point are in accord. See *Fisher*, 133 S. Ct. at 2421 (“In *Grutter*, the Court approved the plan at issue upon concluding that it . . . was limited in time”); *Grutter*, 539 U.S. at 386–87 (Rehnquist, J., dissenting) (collecting sources and declaring, in an opinion joined by Kennedy, that “an important component of strict scrutiny [is] that a program be limited in time”).

244. See Feingold, *supra* note 14, at 264 (proposing a similar stopping point).

245. See generally, e.g., Cohen et al., *Recursive Processes*, *supra* note 31; Harackiewicz et al., *supra* note 31; Miyake et al., *supra* note 31; Sherman et al., *supra* note 31; Walton, Spencer & Erman, *supra* note 14; Yeager et al., *supra* note 94; Yeager et al., *supra* note 96.

246. Walton, Spencer & Erman, *supra* note 14, at 5–6, 10; Yeager & Walton, *supra* note 94, at 268.

where stereotype threat and its reduction is common sense, the need to correct for failures to reduce threat is likely to appear temporary and acceptable.

To explore what types of recalibration the Court might allow, we return to diversity jurisprudence, the touchstone of which is individualized consideration.²⁴⁷ As the Court stated in *Fisher*: “admissions processes [must] ensure that each applicant is evaluated as an individual and not in a way that makes an applicant’s race or ethnicity the defining feature of his or her application.”²⁴⁸ The Court has elaborated that this requires “ensuring that each candidate ‘compete[s] with all other qualified applicants’”²⁴⁹ and has indicated that making specific, identifiable, substantial, determinative score adjustments contravenes the requirement.²⁵⁰

The requirement, the Court reasons, assures that schools’ stated purposes are their actual ones and promotes values that the Justices associate with equal protection. By invalidating laws that do not hew closely to the purposes they purport to serve, the Court claims, it “smoke[s] out illegitimate uses of race,” leaving “little or no possibility that the motive for the classification was illegitimate racial prejudice or stereotype.”²⁵¹ Envisioning diversity to encompass many factors that

247. Siegel, *supra* note 139, at 796–97.

248. *Fisher*, 133 S. Ct. at 2420 (internal quotation marks omitted). See also *Grutter*, 539 U.S. at 337 (“The importance of this individualized consideration in the context of a race-conscious admissions program is paramount.”); *Gratz v. Bollinger*, 539 U.S. 244, 271 (2003) (“Justice Powell’s opinion in *Bakke* emphasized the importance of considering each particular applicant as an individual, assessing all of the qualities that individual possesses, and in turn, evaluating that individual’s ability to contribute to the unique setting of higher education.”); *Parents Involved in Cnty. Sch. v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 798 (2007) (Kennedy, J., concurring in part and concurring in the judgment) (“Race may be one component of . . . diversity [in public school admissions], but other demographic factors, plus special talents and needs, should also be considered.”); *Grutter*, 539 U.S. at 392–93 (Kennedy, J., dissenting) (“There is no constitutional objection to the goal of considering race as one modest factor among many others to achieve diversity, but an educational institution must ensure, through sufficient procedures, that each applicant receives individual consideration and that race does not become a predominant factor in the admissions decisionmaking.”).

249. *Grutter*, 539 U.S. at 335 (quoting *Johnson v. Transp. Agency, Santa Clara Cnty., Cal.* 480 U.S. 616, 638 (1987)).

250. *Gratz*, 539 U.S. at 270–75 (invalidating a policy that “automatically distributes 20 points to every single applicant from an ‘underrepresented minority’ group,” which “has the effect of making ‘the factor of race . . . decisive’ for virtually every minimally qualified underrepresented minority applicant” (quoting *Regents of the Univ. of Cal. v. Bakke*, 438 U.S. 265, 317 (1978))); *Parents Involved*, 551 U.S. at 798 (Kennedy, J., concurring in part and concurring in the judgment) (“What the government is not permitted to do, absent a showing of necessity not made here, is to classify every student on the basis of race and to assign each of them to schools based on that classification.”); *id.* at 723; Siegel, *supra* note 139, at 796–97; Robert C. Post, *The Supreme Court, 2002 Term—Foreword: Fashioning the Legal Constitution: Culture, Courts, and Law*, 117 HARV. L. REV. 4, 71, 73 (2003).

251. *Johnson v. California*, 543 U.S. 499, 506 (2005); *Grutter*, 539 U.S. at 333 (quoting City of

people display in unpredictable combinations, Justice Kennedy and the Court have decided that holistic approaches to selection will better achieve diversity than substantial mechanical race-based score adjustments.

Applying this concern to attempts to recalibrate scores to promote meritocracy is complex because of the Court's simultaneous concern with predictive accuracy and open competition. As already discussed, for the foreseeable future, subsequent performance may be best predicted through standardized group-level recalibration.²⁵² Taking systemic biases into account in flexible—which is generally in this context to say inconsistent—ways would likely worsen predictions of subsequent performance and thus somewhat disserve merit.²⁵³

Yet, recalibrating scores after the fact reduces the extent to which a measure acts as a competition. Conceptions of competition often involve the notion that once the game has been played, the rules cannot be changed in ways that make winners losers and vice versa. But selection is not just a prize. It's an opportunity to pursue further opportunities. Consider a national footrace that uses local time trials to select participants. Ideally, local tracks would all be equal in grade, smoothness, traction, rebound, etc. But if some tracks were clearly faster than others, and if it were impossible to re-run time trials under equal conditions, the racing authority would have three bad options. It could cancel the national race (cancel college), select those with the fastest times regardless of track conditions for the national race (current practice), or adjust times in light of track conditions to select those most likely to win the national race (score recalibration).

The Court has provided contradictory signals as to which approach it would prefer. In *Bakke*, Justice Powell contemplated the possibility of selection criteria that systematically underpredict minorities' subsequent performance and embraced score recalibration as a response. Implicitly

Richmond v. J.A. Croson Co., 488 U.S. 469, 493 (1989) (plurality opinion)); *Parents Involved*, 551 U.S. at 783 (Kennedy, J., concurring in part and concurring in the judgment) (“Absent searching judicial inquiry into the justification for such race-based measures, there is simply no way of determining what classifications are ‘benign’ or ‘remedial’ and what classifications are in fact motivated by illegitimate notions of racial inferiority or simple racial politics.” (quoting *Croson*, 488 U.S. at 493 (plurality opinion))); Rubenfeld, *supra* note 45, at 428, 436–37, 468–69.

252. See *supra* Section III.A.

253. Cf. Jessica Bulman-Pozen, Note, *Grutter at Work: A Title VII Critique of Constitutional Affirmative Action*, 115 YALE L.J. 1408, 1420 (2006) (noting that Title VII jurisprudence tends to see unchecked subjective decision making as inviting bias and so prefers numerical benchmarks); Feingold, *supra* note 14, at 262 (“*Gratz* proves to be of limited value” in evaluating the constitutionality of policies such as mechanical score recalibration); Robinson, *supra* note 187, at 289–91 (observing that diversity gains from using racial classifications will rarely justify their use given the increased harm to innocent third parties that accompany such use).

presuming that unbiased, similarly predictive measures might not be available, Powell faced a choice like that of the national racing authority. He assumed that the school would select a class—that the race would go forward—and framed the choice as between using recalibrated or biased scores. The former, he indicated, would be acceptable, likely raising no equal protection concerns at all.

The Court's more recent decision in *Ricci*, by contrast, stressed the importance of honoring the results of competitive exams. As already discussed, because the Court focused on the content of the exam to the relative exclusion of consideration of its context, the Court perceived New Haven to have taken effective steps to ensure that its test was not racially biased. After noting the statutory bar on adjusting scores on employment tests, the Court then framed a choice between abandoning and embracing test results it deemed unbiased. Only the latter, it announced, was permissible: “once [the selection] process has been established and employers have made clear their selection criteria, they may not then invalidate the test results, thus upsetting an employee’s legitimate expectation not to be judged on the basis of race.”²⁵⁴ Although the Court did not address the situation either of a test that it recognized as infected by bias or of an organization that announces a plan to recalibrate scores prior to the test’s administration, its language suggests that it would reach a similar result.²⁵⁵

To reframe *Ricci* in terms of the racing analogy, it was as though a national authority barred from adjusting racing times had ensured that all local tracks were equal and then faced a choice between whether to cancel the national race or let the runners with the best local times run.²⁵⁶ The harder and more pertinent question arises if the authority’s efforts to

254. *Ricci v. DeStefano*, 557 U.S. 557, 585 (2009).

255. Announcing plans to recalibrate scores at the outset would reduce reliance interests. But it would also increase the visibility and salience of the racial classifications, an outcome the Court has elsewhere deemed fatally objectionable. *See infra* notes 258–272 and accompanying text.

256. A racing analogy also helps answer a different objection: if stereotyped students’ scores are to be raised when they come to underpredict their subsequent performance, should they be lowered when they overpredict their subsequent performance? *Cf. Grutter*, 539 U.S. at 350 (Thomas, J., dissenting) (“No one would argue that a university could set up a lower general admissions standard and then impose heightened requirements only on black applicants. Similarly, a university may not maintain a high admissions standard and grant exemptions to favored races.”). Imagine a national race that uses local time trials of varying quality to select participants for a national competition. Between the two rounds, storms damage all tracks but inferior tracks most. Thus, runners on initially inferior tracks who barely beat those on superior tracks in round one would likely lose to such opponents in round two. If one focused solely on subsequent performance, one would select for round two slightly slower round-one runners with access to superior tracks. But that would be inconsistent with most notions of open competition.

equalize local tracks failed—they might be the same length (same test content) but vary in quality (some students might labor under stereotype threat while others do not). Should the authority cancel the national race entirely or go ahead despite the potential illegitimacy of holding a national race in which the fastest runners may have little meaningful opportunity to secure a place? At least in its language, the *Ricci* Court indicated its preference for the latter.

The Court also requires individualized consideration in the diversity context because it perceives doing so to promote additional values that the Court associates with equal protection.²⁵⁷ Those include safeguarding against racial politics, essentialism, and depreciation and ensuring individual dignity and autonomy. The relationship of each value to score recalibration is complex.

Governmental action based on race is dangerous in part, the Court has asserted, because “appearances do matter.”²⁵⁸ Official policies that visibly depend upon racial criteria communicate that race is relevant,²⁵⁹ and, according to the Court, thus “may balkanize us into competing racial factions.”²⁶⁰ Five current Justices have reasoned that “race-based reasoning and the conception of a Nation divided into racial blocs” threatens an “escalation of racial . . . conflict” and “a politics of racial hostility.”²⁶¹ Kennedy added in *Parents Involved* that such policies make “race . . . a bargaining chip in the political process” and thereby breed “new divisiveness.”²⁶²

257. Rubenfeld, *supra* note 45, at 428–29, 436–40, *passim* (discussing narrow tailoring as cost-benefit analysis); Fallon, *supra* note 45, at 1271, 1307–09 (describing strict scrutiny as in some cases seeking a rights-skewed balance between marginal losses of rights and marginal gains in governmental interests).

258. *Shaw v. Reno*, 509 U.S. 630, 647 (1993); Siegel, *supra* note 131, at 1300–01; Ayres & Foster, *supra* note 45. Cf. *McCleskey v. Kemp*, 481 U.S. 279, 286–87, 291 (1987) (rejecting equal-protection challenge to a death sentence imposed by a jury in the individualized context of a single prosecution notwithstanding statistics revealing that the state used juries in ways that produced stark state-level racial disparities in death-penalty sentencing).

259. Siegel, *supra* note 139, at 809–11.

260. *Shaw*, 509 U.S. at 657. See also *Schuette v. Coal. to Defend Affirmative Action*, No. 12-682, slip op. at 12 (S. Ct. Apr. 22, 2014) (plurality opinion) (similar).

261. *Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 746 (2007) (plurality opinion) (quoting *Metro Broad., Inc. v. FCC*, 497 U.S. 547, 603 (1990) (O’Connor, J., dissenting)); *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 493 (1989) (plurality opinion); *id.* at 746 (plurality opinion); *Ricci v. DeStefano*, 557 U.S. 557, 583–84, *passim* (2009) (implying that making decisions based on race rather than excellence could breed hostility among whites who perceive themselves to have been incorrectly deemed unworthy of a benefit); Siegel, *supra* note 131, at 1294–95, 1298–99; Siegel, *supra* note 139, at 806–07, 836.

262. *Parents Involved*, 551 U.S. at 797 (Kennedy, J., concurring in part and concurring in the judgment). Jurists have hypothesized that balkanization could result if governmental programs

Race-based governmental policies, the Court has also stated, send messages that threaten to reinforce racial essentialism and depreciation: They imply that all members of a racial group think alike, and they “promote notions of racial inferiority.”²⁶³

When officials draw racial lines, the Court has insisted, they also potentially intrude upon individuals’ constitutional interests in dignity and autonomy.²⁶⁴ Justice Kennedy made the case particularly forcefully in

benefitting underrepresented minorities cause minorities to internalize a sense of entitlement; breed hostility among whites who perceive themselves to have been deemed unworthy or otherwise denied a benefit because of race; spark cross-racial tension; or provide tools to those pursuing racial politics. *See, e.g.*, *Ricci*, 557 U.S. at 604–05 (Alito, J., concurring); Siegel, *supra* note 131, at 1294–95, 1298–99; Siegel, *supra* note 139, at 806–07, 836.

263. *Grutter v. Bollinger*, 539 U.S. 306, 328 (2003) (quoting *Croson*, 488 U.S. at 493 (plurality opinion)). *See also Shaw*, 509 U.S. at 643 (“[E]ven in the pursuit of remedial objections, an explicit policy of assignment by race may serve to stimulate our society’s latent race consciousness, suggesting the utility and propriety of basing decisions on a factor that ideally bears no relationship to an individual’s worth or needs” (alteration in original) (quoting United Jewish Orgs. of Williamsburgh, Inc. v. Carey, 430 U.S. 144, 173 (1977) (Brennan, J., concurring in part))); *id.* at 647 (asserting that race-based voting districts “reinforce[] the perception that members of the same racial group—regardless of their [differences]—think alike, share the same political interests, and will prefer the same candidates at the polls”); *id.* at 657; *Parents Involved*, 551 U.S. at 782 (Kennedy, J., concurring in part and concurring in the judgment); *id.* at 746 (plurality opinion); *Adarand Constructors, Inc. v. Pena*, 515 U.S. 200, 229 (1995) (claiming that any racial preference “inevitably is perceived by many as resting on an assumption that those who are granted this special preference are less qualified in some respect” (quoting *Fullilove v. Klutznick*, 448 U.S. 448, 545 (1980) (Stevens, J., dissenting))); *Croson*, 488 U.S. at 516–17 (Stevens, J., concurring in part and concurring in the judgment) (criticizing racial preference for imposing “stigma on its supposed beneficiaries”); *Johnson v. California*, 543 U.S. 499, 507 (2005); *Coalition to Defend Affirmative Action*, slip op. at 11–12 (plurality opinion) (reiterating concern with “demeaning stereotypes” and the notion that all members of a race think alike); Siegel, *supra* note 139, at 836.

Research into stereotype threat has confirmed that racial classifications that raise the salience of negative racial stereotypes can have effects like those Kennedy describes on members of negatively stereotyped groups. *See, e.g.*, Steele & Aronson, *supra* note 13; Danaher & Crandall, *supra* note 24. But racial classifications do not always have this effect. In some cases, they facilitate interventions to reduce stereotype threat and undo harms like those Kennedy describes. *See Marx & Roman*, *supra* note 115 (“[F]emale role models can buffer women’s math test performance”); Stout et al., *supra* note 115, at 255 (“[E]xposure to female STEM experts promoted positive implicit attitudes and stronger implicit identification with STEM, greater self-efficacy in STEM, and more effort on STEM tests” among women). Cf. Purdie-Vaughns et al., *supra* note 116 (discussing the impact of working in a racially diverse workplace); McIntyre, Paulson & Lord, *supra* note 115, at 83 (discussing the impact of “reminding women of other women’s achievements”).

264. *See supra* note 175 and accompanying text. Cf. *Fisher v. Univ. of Tex. at Austin*, 133 S. Ct. 2411, 2418–19 (2013) (“[B]ecause racial characteristics so seldom provide a relevant basis for disparate treatment, the Equal Protection Clause demands that racial classifications . . . be subjected to the most rigid scrutiny.” (alterations in original) (citations and internal quotation marks omitted)). On the authority to be attributed to points of agreement in the plurality opinion and determinative concurrence in *Parents Involved*, which together garnered the assent of all five Justices in the majority, consider *Marks v. United States*, 430 U.S. 188, 193 (1977) (“When a fragmented Court decides a case and no single rationale explaining the result enjoys the assent of five Justices, the holding of the Court may be

Parents Involved: “Under our Constitution the individual, child or adult, can find his own identity, can define her own persona . . .”²⁶⁵ But a governmental racial classification imposes a “state-mandated racial label” that the “individual is” “forced to live under.”²⁶⁶ Doing so “is inconsistent with the dignity of individuals in our society” and threatens “to reduce children to racial chits valued and traded according to one school’s supply and another’s demand.”²⁶⁷

In the diversity context, the Court has reasoned, viewing race as one trait—with no preordained value—among many tends to make its importance less visible. That, in turn, can, in Justice Kennedy’s words, dampen its potentially “corrosive” effect on “discourse.”²⁶⁸ By contrast, Kennedy has argued, schools that use race as shorthand for diversity replace holistic evaluations of individuals with racial labels that deny each person autonomy over self-definition and dignity.²⁶⁹

Meritocratic score recalibration along racial lines operates differently, promoting—but also sometimes modestly contravening—equal protection values. Rather than communicate a broadly essentializing racial message that all minorities think and behave similarly along many axes, they rest on a narrow intraracial similarity: increased susceptibility to specific kinds of stereotype threat. Otherwise, such policies presume, members of particular

viewed as that position taken by those Members who concurred in the judgments on the narrowest grounds.” (internal quotation marks omitted)).

265. *Parents Involved*, 551 U.S. at 797 (Kennedy, J., concurring).

266. *Id.* See also Ian Ayres, *Narrow Tailoring*, 43 UCLA L. REV. 1781, 1797 (1996).

267. *Parents Involved*, 551 U.S. at 797–98 (Kennedy, J., concurring); *Coalition to Defend Affirmative Action*, slip op. at 16 (plurality opinion) (describing an equal-protection aspiration of “equal dignity”). But cf. *Parents Involved*, 551 U.S. at 795 (acknowledging and then rejecting as a basis for action that “[f]rom the standpoint of the victim, it is true, an injury stemming from racial prejudice can hurt as much when the demeaning treatment based on race identity stems from bias masked deep within the social order as when it is imposed by law”).

268. *Parents Involved*, 551 U.S. at 797.

269. See *supra* notes 264–267. See also *Metro Broad., Inc. v. FCC*, 497 U.S. 547, 632 (1990) (Kennedy, J., dissenting) (rejecting the view that it “is worth the cost of discriminating among citizens on the basis of race because it will increase the listening pleasure of media audiences”). Cf. *Brown v. Bd. of Educ. of Topeka*, 347 U.S. 483, 493 (1954) (overturning doctrine of separate but equal because it “deprive[s] the children of the minority group of equal educational opportunities”); *Rice v. Cayetano*, 528 U.S. 495, 523 (2000) (“Race cannot qualify some and disqualify others from full participation in our democracy. All citizens, regardless of race, have an interest in selecting officials who make policies on their behalf, even if those policies will affect some groups more than others.”); *id.* at 517 (“An inquiry into ancestral lines is not consistent with respect based on the unique personality each of us possesses . . .”); *Lawrence v. Texas*, 539 U.S. 558, 574 (2003) (“These matters, involving the most intimate and personal choices a person may make in a lifetime, choices central to personal dignity and autonomy, are central to the liberty protected by the Fourteenth Amendment.” (internal quotation marks omitted))).

racial groups should not be treated as a class apart. Recalibrating the scores of racial minority students who vary in their academic qualifications will maintain those intragroup distinctions: A student with a higher assessed qualification will still be higher after accounting for the effects of stereotype threat. Similarly, the core insight behind score recalibration is that because stereotyped individuals often face unique performance-depressing barriers—a group-based liability—they also often have substantial unrecognized ability. And because score recalibration aims to select those candidates likely to best achieve, some commentators generally opposed to affirmative action advocate for recalibrating scores to better predict subsequent performance.²⁷⁰ If nonstereotyped individuals nonetheless resent resultant losses of benefits they do not merit—which some surely will—their dissatisfaction may reach the limits of the Court’s indulgence.²⁷¹ Finally, because score recalibration uses race to facilitate consideration of people’s potential, it promotes equal protection values of dignity and autonomy. In this context, using individualized review in selection as a way to obscure the role of racial classifications in the process would also hinder consideration of people based on their true characteristics. The way to stop selection on the basis of race is to start classifying on the basis of race—and mechanically so.

Taken together, cross-cutting effects on equal protection values provide modest guidance in choosing between mechanical score

270. RICHARD J. HERRNSTEIN & CHARLES MURRAY, THE BELL CURVE: INTELLIGENCE AND CLASS STRUCTURE IN AMERICAN LIFE 280 (1994) (“If the SAT is biased against blacks, it will *underpredict* their college performance. . . . It would be as if the test underestimated the ‘true’ SAT score of the blacks, so the natural remedy for this kind of bias would be to compensate the black applicants by, for example, adding the appropriate number of points onto their scores.”); Paul R. Sackett & Steffanie L. Wilk, *Within-Group Norming and Other Forms of Score Adjustment in Preemployment Testing*, 49 AM. PSYCHOLOGIST 929, 933 (1994) (calling “score adjustment[s]” “a technically appropriate solution” in response to a finding of latent ability). See also Leanne S. Son Hing, D. Ramona Bobocel & Mark P. Zanna, *Meritocracy and Opposition to Affirmative Action: Making Concessions in the Face of Discrimination*, 83 J. PERSONALITY & SOC. PSYCHOL. 493, 505 (2002) (finding that views toward affirmative action programs improve in response to perceived workplace discrimination); Walton, Spencer & Erman, *supra* note 14, at 25 (noting that the use of score corrections for the purpose of merit mitigates concerns related to whether affirmative action is harmful to stereotyped students or reinforces stereotypes).

271. Primus, *supra* note 40, at 570 (arguing that although enormous balkanization followed *Brown v. Board of Education*, courts correctly refrained from allowing “such hostility to veto government action aimed at improving the position of disadvantaged groups”); Walton, Spencer & Erman, *supra* note 14, at 27. See also Feingold, *supra* note 14, at 259 (deeming policies akin to score recalibration “immune to the . . . ‘innocent third party’ objections”). The Court may be especially unsympathetic because forbidding recalibration could cause resentment among members of stereotyped groups who would know that people like them would predictably be denied opportunities in favor of less meritorious nonstereotyped individuals.

recalibration and individualized review. If Justices were to perceive score recalibration, on balance, to promote equal protection values, they might be inclined to permit it to operate in the highly visible way that they associate with mechanical approaches. That outcome, however, seems unlikely given the Court's unrelenting hostility toward mechanical score adjustments. The alternative view—that score recalibration is more threat than boon to equal protection values—would likely lead the Court to require individualized consideration. That is especially so because the benefits of mechanical score recalibration can be realized—albeit to lesser degrees—through individualized selection.

A school considering whether to implement meritocratic score recalibration faces bad choices. Taking no action means accepting weaker students and exacerbating racial inequality. Yet, ambiguities and tensions in the Court's equal protection jurisprudence make it uncertain whether the Court would permit any form of score recalibration. The Court would likely bar the most accurate approach, mechanical score recalibration. Somewhat less risky and less effective would be taking stereotype threat into account as part of an individualized selection process.²⁷² In any case, the risk of mispredicting how the Court will apply its indeterminate jurisprudence falls on schools.

C. DISPARATE EFFECTS

Compounding the dilemma for schools deciding whether to implement score recalibration is the possibility that Title VI regulations may sometimes require them to do so. Recall their mandate: no disparate racial effects that lack a substantial, legitimate justification or can be avoided through an equally effective, less-discriminatory alternative. Application of the first two criteria is straightforward. Schools that use measures infected by stereotype threat in selection produce racially disparate effects.²⁷³ And

272. See Walton, Spencer & Erman, *supra* note 14, at 27–28 (explaining how a school could “educate selection officers of the bias in specific performance measures and allow them to weigh this information in making individualized evaluations of candidates . . . on the merits of their entire applications”).

273. The requirement of disparate group-level effects distinguishes a disparate impact claim from an accommodation claim, which can rest instead on the harm to a single individual that flows from the interaction between a given policy and the particular limitations of that person. See, e.g., 42 U.S.C. § 12112(b) (defining disability discrimination in employment to include “(5)(A) not making reasonable accommodations to the . . . limitations of an otherwise qualified individual”); Trans World Airlines, Inc. v. Hardison, 432 U.S. 63 (1977) (similar for religious accommodations); Henrietta D. v. Bloomberg, 331 F.3d 261, 273–77 (2d Cir. 2003) (finding that disability accommodation claim does not require “that plaintiffs identify a comparison class of similarly situated individuals given preferential treatment” and collecting cases reaching same result (internal quotation marks omitted)). This aspect of disparate

they have a substantial justification: meritocracy. In some cases, schools can at reasonable cost and without loss of predictive accuracy eliminate threat from the infected measure, switch to measures uninfected by threat, or limit the use of the infected measure. Under Title VI regulations, they must do what they should do in any case and act. Where these responses do not solve the problem, the question becomes whether performance-maximizing score recalibration forms an equally effective alternative to use of the infected measure.²⁷⁴ This section argues that they do, and that if they are otherwise constitutional they would be required.

Schools that use measures whose predictive accuracy stereotype threat reduces will generally find that at least some form of meritocratic score recalibration constitutes an equally effective alternative. Recalibrating scores as part of an individualized selection process would generally qualify for schools already engaging in individualized review of candidates. There, the change would impose little additional cost and increase predictive accuracy. Where a school has used automatic algorithms to admit students, by contrast, individualized review could bring soaring administrative costs that would render it less than equally effective. Mechanical score recalibration would likely be an equally effective alternative regardless of whether a school already engages in individualized review. It offers a cheap and precise alternative to inaction. Contrary to frequent judicial and popular intuitions that reducing racial disparities in outcomes comes at the expense of meritocracy, score recalibration would select candidates more likely to perform better. Title VI regulations likely require that schools unable to reduce unjustified racial disparities in other ways implement score recalibration.

Schools could thus find themselves caught between the strictures of Title VI disparate-effect regulations and statutory and constitutional bans on racial classifications. That tight spot would bring to the fore the potential conflict that *Ricci v. DeStefano* sidestepped between sub-constitutional disparate effects tests and constitutional equal protection

impact law is in tune with how stereotype threat harms entire groups. Researchers sometimes analogize the effects of negative stereotypes to a “chilly climate.” Walton et al., *supra* note 30. Although there are various ways to remedy a cold room—from raising the heat to providing sweaters—some individuals are also more susceptible to cold than others. Similarly, stereotype threat primarily arises from a negative social circumstance—the existence of pervasive negative intellectual stereotypes. This fact harms everyone to whom the stereotype applies even as some individuals may be somewhat more susceptible to this harm than others.

274. Potential-maximizing score recalibration would not be equally effective alternatives to performance-maximizing score recalibration because it would less well predict how applicants would perform if selected.

requirements.²⁷⁵ *Ricci* involved a parallel Title VII provision barring racial disparities in employment, which the Court construed after New Haven threw out the results of a test on which white firefighters had scored disproportionately well.²⁷⁶ The holding of the case was statutory: that Title VII only permits employers to engage in intentional discrimination if they have a “strong basis in evidence” for believing that they will otherwise be subject to liability for producing illegal disparate racial outcomes.²⁷⁷ But as others have observed, the cadences of the majority opinion were constitutional.²⁷⁸ And with the majority having expressly reserved the question of the constitutionality of the Title VII bar on producing certain racial disparities,²⁷⁹ many have read its analysis of the statutory question to presage its approach to the constitutional one.²⁸⁰

But there are several ways to read *Ricci*. One moral from *Ricci* may be that judges and not employers should have the final say on what steps are necessary to avoid disparate-impact liability.²⁸¹ Here, the strong-basis-in-evidence test would serve to ensure that judges scrutinize any employer claim that intentional discrimination was necessary to avoid illegal racial

275. If Title VI disparate effects regulations violate the Constitution, the ones whose rights are violated are those whose outcomes factor into the disparate-effects analysis—the test-takers rather than the university seeking to utilize the test results. Under the third-party standing rule, a party “generally must assert his own legal rights and interests, and cannot rest his claim to relief on the legal rights or interests of third parties.” *Kowalski v. Tesmer*, 543 U.S. 125, 129 (2004) (internal quotation marks omitted). While an exception exists where a third party has a close relationship to one hindered in exercising her own rights, *e.g.*, *id.* at 129–30, a surer route to a constitutional ruling on the merits would be a challenge by an applicant who received no additional points to her test score while applicants of different races did.

276. The Title VII bar on bottom-line analysis, by contrast, does not endanger score recalibration. In *Connecticut v. Teal*, 457 U.S. 440 (1982), the Court confronted a promotion exam purported to violate the Title VII bar on non-job-related employment practices that produce racial disparities. *Id.* at 443–44. After being sued, the employer promoted those black employees who passed the test at a higher rate than similarly situated white employees. *Id.* at 444. As a result, blacks were promoted at a higher rate overall than whites. *Id.* The Court found this bottom line irrelevant, insisting instead that the focus must be on each employment practice. *Id.* at 452–56. By contrast, in our case—concerning the use of biased measures with a substantial, legitimate justification—there is no freestanding practice that violates a bar on inadequately justified disparate racial effects. The biased measures only become potentially illegal when score corrections constitute an equally effective alternative. Absent score recalibration, that would not be the case. But in our case, score recalibration is the only potentially equally effective alternative. Its existence creates the possibility of liability rather than offering a defense to it.

277. *Ricci v. DeStefano*, 557 U.S. 557, 585, 577 (2009).

278. Primus, *supra* note 65, at 1344, 1356–62; Siegel, *supra* note 9, at 53–54.

279. *Ricci*, 557 U.S. at 584.

280. *E.g.*, Primus, *supra* note 65; Siegel, *supra* note 131; Norton, *supra* note 186, at 229, 236, 239, 242–44; Adams, *supra* note 138, at 839.

281. Primus, *supra* note 65, at 1344–45, 1368–69; Rich, *supra* note 34, at 76.

disparities.²⁸² If so, one would expect the test to apply to Title VI and associated regulations, which similarly bar intentional discrimination and some disparate effects.²⁸³ Race-based score recalibration aimed at maximizing performance would generally meet this test with little difficulty, for it is often a cheap, effective, evidence-based way to less often choose weaker students of one race over stronger students of another.²⁸⁴

Another way to view *Ricci* is as a case of an employer with an illicit purpose. Here, the starting point is that the Court found New Haven to have engaged in intentional discrimination,²⁸⁵ which requires a racial classification or a discriminatory purpose. And because New Haven classified a test rather than people, it made no racial classification.²⁸⁶ Thus, New Haven must have had a discriminatory purpose.²⁸⁷ But the Court did not identify which of New Haven's purposes was illegal. Possibilities include New Haven's attempt to seek racial equality and integration, the race consciousness the city displayed as a necessary consequence of seeking to comply with disparate-impact requirements, or the heightened race consciousness that New Haven's particular approach to avoiding disparate-impact liability involved.²⁸⁸ If the Court intended to equate either of the first two possibilities with a discriminatory purpose, then the Title VI

282. Primus, *supra* note 65, at 1344–45, 1368–69; Rich, *supra* note 34, at 76.

283. See *supra* note 278 and accompanying text.

284. Schools will, of course, want to ensure that they have a strong basis in evidence for the particular score recalibration that they implement. Hence, better not to add a set amount of points to the scores of all stereotyped minorities—or to exclude all Asians from score recalibration if some Asian ethnicities do tend to be subject to stereotype threat—if the research shows that groups vary in terms of stereotype-threat effects. See *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 506 (1989) (criticizing an ostensibly remedial affirmative-action program that benefited “Spanish-speaking, Oriental, Indian, Eskimo, [and] Aleut persons” despite there having been “absolutely no evidence of past discrimination” against members of those groups); *Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 723–34 (plurality opinion) (criticizing two diversity plans for “viewing race exclusively in white/nonwhite . . . and black/‘other’ terms”).

285. *Ricci*, 557 U.S. at 579–80 (concluding that New Haven “rejected the test results solely because the higher scoring candidates were white”).

286. Rich, *supra* note 34, at 74–76; Siegel, *supra* note 131, at 1325.

287. Adams, *supra* note 138; Norton, *supra* note 186, at 229; Primus, *supra* note 65, at 1352–53; *Ricci*, 557 U.S., at 594–96 (Scalia, J., concurring); John V. Wintermute, Comment *Remedying Race-Based Decision-Making: Reclaiming the Remedial Focus of Affirmative Action after Fisher v. University of Texas at Austin*, 44 SETON HALL L. REV. 557, 563–73 (2014).

288. *Ricci*, 557 U.S. at 594–96 (Scalia, J., concurring); Adams, *supra* note 138; Norton, *supra* note 186, at 229; Primus, *supra* note 65, at 1344–45, 1370–74; Brodin, *supra* note 186, at 230–31; Rich, *supra* note 34, at 76; Siegel, *supra* note 131, at 1325; Siegel, *supra* note 9, at 54, 61. Cf. *Schuette v. Coal. to Defend Affirmative Action*, No. 12-682, slip op. at 46 (S. Ct. Apr. 22, 2014) (Sotomayor, J., dissenting) (“In my colleagues’ view, examining the racial impact of legislation only perpetuates racial discrimination. . . . It is this view that works harm, by perpetuating the facile notion that what makes race matter is acknowledging the simple truth that race *does* matter.”).

disparate-effects test likely violates the Constitution. But that probably was not the intent of *Ricci*'s author.²⁸⁹ In addition to writing *Ricci*, Justice Kennedy penned *Fisher v. University of Texas* and the decisive concurrence in *Parents Involved*, both of which tolerate race-conscious activities and the latter of which commends racial integration as a governmental motive. In *Parents Involved*, Kennedy reassured that school boards "may pursue the goal of bringing together students of diverse backgrounds and races through . . . means" that "are race conscious but do not lead to different treatment based on a [racial] classification."²⁹⁰ Examples, he wrote, included "strategic site selection of new schools" and "drawing attendance zones with general recognition of the demographics of neighborhoods."²⁹¹ In *Fisher*, Kennedy's majority opinion similarly stressed the value—and implicitly the acceptability—of nonclassificatory approaches to racial and other diversity.²⁹² It is more likely that New

289. See Siegel, *supra* note 131, at 1326–28 ("If *that* is what *Ricci* holds, then Justice Kennedy would seem to be suggesting that a government's decision to select a facially neutral policy that promotes employee diversity—the very sort of decision he went out of his way to affirm in *Parents Involved*—violates Title VII and possibly the Equal Protection Clause as well."). But cf. Siegel, *supra* note 9, at 58 & n.287 (noting, before the case settled and was dismissed, that "[w]ith conservative interest in challenging disparate impact on . . . constitutional grounds high, the Court ha[d] taken another disparate impact case for the 2013 Term," *Mt. Holly Gardens Citizens in Action, Inc. v. Twp. of Mount Holly*, 658 F.3d 375 (3d Cir. 2011), cert. granted, 133 S. Ct. 2824 (2013), cert. dismissed, 134 S. Ct. 636 (2013)).

290. *Parents Involved in Cmty. Sch. v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 789 (2007) (Kennedy, J., concurring in part and concurring in the judgment).

291. *Id.*

292. *Fisher v. Univ. of Texas at Austin*, 133 S. Ct. 2411, 2420 (2013). See also Jack Balkin, *Why Fisher is Important*, BALKINIZATION (June 24, 2013), <http://balkin.blogspot.com/2013/06/why-fisher-is-important.html> ("According to seven Justices in *Fisher*, [a facially race-neutral program] appears to be perfectly permissible, because it is the constitutionally preferable alternative to using race directly as a factor in admissions."); Rich, *supra* note 9, at 1574–79 ("The very same affirmative action decisions that apply strict scrutiny to formally race-based affirmative action nevertheless suggest that strict scrutiny would not constrain facially neutral attempts to pursue similarly race conscious objectives."). It bears emphasizing that readings of *Ricci* that Justice Kennedy would be unlikely to embrace nonetheless remain available to a future Court, which could choose to accord Kennedy's concurrence in *Parents Involved* relatively little weight notwithstanding the *Marks* rule described *supra* note 264. See *Nichols v. United States*, 511 U.S. 738, 745 (1994) ("[The *Marks*] test is more easily stated than applied . . ."); *United States v. Johnson*, 467 F.3d 56, 65–66 (1st Cir. 2006) ("[T]he Supreme Court itself has moved away from the *Marks* formula."); Melissa M. Berry, Donald J. Kochan & Matthew Parlow, *Much Ado About Pluralities: Pride and Precedent Amidst the Cacophony of Concurrences, and Re-Percolation after Rapanos*, 15 VA. J. SOC. POL'Y & L. 299, 331 (2008) (noting "the Court's own failure to apply the *Marks* doctrine regularly"); Adam S. Hochschild, *The Modern Problem of Supreme Court Plurality Decision: Interpretation in Historical Perspective*, 4 WASH. U. J.L. & POL'Y 261, 282–83 (2000) (citing cases in which the Court has not applied *Marks* when it arguably should have done so); Saul Levmore, *Ruling Majorities and Reasoning Pluralities*, 3 THEORETICAL INQUIRIES L. 87, 109 & n.39 (2002) (citing "some evidence" of the possibility that "judges are content to leave things open for future misinterpretation on the grounds that future courts will regard these narrowest-majority precedents as relatively weak"); Justin F. Marceau, *Lifting the Haze of Baze: Lethal Injections, the*

Haven erred in how it sought to comply with disparate impact than in deciding to comply at all.

There are good reasons to think that *Ricci* is best understood as a case about New Haven's approach to avoiding liability for racial disparities. After all, the Court criticized at length New Haven's particular remedy: throwing out the results of an apparently fair and effective test on unambiguously, highly visible racial grounds; and stripping a small number of identifiable, innocent white firefighters of benefits they had earned according to preexisting rules through hard work, economic investment in test preparation, and superior performance.²⁹³ On this view, the statutory holding saved the Court from making explicit its core concern—that it was New Haven's means of complying with Title VII that violated equal protection. That logic suggests that compliance with the Title VII bar on certain racial disparities would never be a winning defense against an otherwise valid equal protection claim.²⁹⁴ But it also suggests that *Ricci*

Eighth Amendment, and Plurality Opinions, 41 ARIZ. ST. L.J. 159 (2009); Mark Alan Thurmon, Note, *When the Court Divides: Reconsidering the Precedential Value of Supreme Court Plurality Decisions*, 42 DUKE L.J. 419, 438–42 (1992); Linas E. Ledebur, Comment, *Plurality Rule: Concurring Opinions and a Divided Supreme Court*, 113 PENN ST. L. REV. 899 (2009). Cf. *Grutter v. Bollinger*, 539 U.S. 306, 325 (2003) (declining to apply the *Marks* rule when accepting the conclusion of a swing Justice in a prior case). One response to the argument that *Ricci* must be read in light of *Parents Involved* and *Fisher* is that the cases are distinguishable. On this view, *Parents Involved* and *Fisher* are peculiar to the education context, which Justice Kennedy perceives as playing a unique role in constructing the U.S. civic sphere. See Gerken, *supra* note 140. For purposes of this Article, which only concerns education, the distinction makes little difference.

293. Adams, *supra* note 138, at 842; Primus, *supra* note 65, at 1345, 1370–74; Rich, *supra* note 34, at 75–77; Siegel, *supra* note 131, at 1285–86, 1325–34. On the possibility that both black and white firefighters sought a fair process in which race would not be determinative and on the failure of New Haven to take steps, like testing the validity of its test, that might have addressed such concerns no matter their result, see Siegel, *supra* note 131, at 1342–45; Norton, *supra* note 186, at 257.

294. That federal regulations require what the Constitution and statute might forbid is unlikely to alter the analysis. While federal officials might have a compelling interest in complying with the effects test so long as its constitutionality is in doubt, that possibility “does not bear on whether Congress violated the Constitution by passing the statute.” See Primus, *supra* note 65, at 1379–82 (writing that though “seven of the now-sitting Justices have endorsed the idea of compliance with federal law as a compelling interest,” a Title VII provision analogous to the effects test could only survive judicial review if it were narrowly tailored to some other compelling interest); *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 500 (1989) (indicating that it is permissible to remedy prior discrimination caused by an identifiable “constitutional or statutory violation” so long as there is a “strong basis in evidence for [the] conclusion that remedial action was necessary” (internal quotation marks omitted)), cited with approval in *Ricci v. DeStefano*, 557 U.S. 557, 582 (2009). And although Section 5 of the Fourteenth Amendment vests Congress with the “power to enforce, by appropriate legislation, the provisions” of the Amendment, U.S. CONST. amend. XIV, § 5, which some have seen as an alternate path to constitutionality for similar statutes, see Primus, *supra* note 40, at 495 n.4 (collecting sources), the Court today reads that provision narrowly, see Kenji Yoshino, *The New Equal Protection*, 124 HARV. L. REV. 747, 768–73 (2011) (arguing that the Court has increasingly limited congressional Enforcement Clause powers). Cf. *Regents of University of California v. Bakke*, 438 U.S. 265, 307 (Powell, J.)

does not doom all prohibitions on racial disparities. If so, the disparate-effects test would survive to bar racial disparities to the extent that equally effective alternatives existed that were consistent with equal protection.²⁹⁵ The question from Part III.B of whether race-based score recalibration could survive strict scrutiny would be the flipside of the question whether Title VI regulations required them. If not forbidden, it would be compelled.

Schools deciding whether and how to implement score recalibration face a murky legal landscape. Mechanical score recalibration may violate equal protection. If not, inaction could violate Title VI regulations. Score recalibration via individualized selection provides a potential middle ground: more likely permissible under equal protection than mechanical score recalibration, yet still capable of eliminating the portion of racial effects that could violate Title VI regulations. To avoid legal risks altogether, a school would have to abandon threat-infected measures, which could mean decoupling selection from merit.

CONCLUSION

The use of affirmative meritocracy in education—policies and practices to release and recognize ability that is otherwise hidden by stereotype threat—has the potential to transform antidiscrimination law, methods, and outcomes. Take strict scrutiny. A key form of affirmative

(concluding that the “governmental interest in preferring members of the injured groups at the expense of others is substantial” following “administrative findings of constitutional or statutory violations”); Michael Coenen, *Constitutional Privileging*, 99 VA. L. REV. 683 (2013) (noting the judicial tendency to extend constitutional claims preferential procedural and remedial treatment).

Because many agencies enforce Title VI, none is likely to receive *Chevron* deference in interpreting it. See, e.g., Sapna Kumar, *Expert Court, Expert Agency*, 44 U.C. DAVIS L. REV. 1547, 1570 (“[I]f two agencies administer the same statute, neither is eligible for *Chevron* deference.”). The Court also well might deem Title VI to be unambiguously coextensive with constitutional equal protection, Alexander v. Sandoval, 532 U.S. 275, 280–81 (2001) (deeming it “beyond dispute” that Title VI “prohibits only intentional discrimination”); Gratz v. Bollinger, 539 U.S. 244, 276 n.23 (citing *Sandoval* and other cases as establishing “that discrimination that violates the Equal Protection Clause of the Fourteenth Amendment committed by an institution that accepts federal funds also constitutes a violation of Title VI”), which would render agency regulations to the contrary powerless to alter application of the statute, Nat’l Cable & Telecom. Ass’n v. Brand X Internet Svcs., 545 U.S. 967, 982 (2005) (“[A] judicial precedent holding that the statute unambiguously forecloses the agency’s interpretation . . . displaces a conflicting agency construction.”).

295. Primus, *supra* note 65, at 1345, 1370–74; Norton, *supra* note 186, at 236, 243–44; Adams, *supra* note 138, at 842, 862, 871; Roberto L. Corrada, Ricci’s *Dicta: Signaling a New Standard for Affirmative Action Under Title VII?* 46 WAKE FOREST L. REV. 241, 256 (2011); Harris & West-Faulcon, *supra* note 186, at 84; Rich, *supra* note 34, at 75–77; Michael Selmi, *Understanding Discrimination in a “Post-Racial” World*, 32 CARDOZO L. REV. 833, 846 (2011); Siegel, *supra* note 131, at 1285–86, 1331–45; Rich, *supra* note 9 (arguing that the Court subjects what he terms “inferred classifications” to strict scrutiny in addition to racial classifications and discriminatory purposes).

meritocracy—race-based score recalibration—illuminates uncertainties concerning its constituent ends and means tests. Meritocratic race-based score recalibration that aims to best predict subsequent performance—and thus does not target past wrongs or diversity—raises the open question whether the merit that it pursues is a compelling governmental interest. We argue that it probably is.

If so, race-based score recalibration would raise ambiguities concerning what the Court means by merit and narrow tailoring. The Court has portrayed meritocratic selection as resting on both likely subsequent performance and open competition. Post-hoc mechanical race-based score recalibration maximizes the former value but sits less easily with the latter. Other equal protection values that the Court recognizes also bear uncertain relationships to race-based score recalibration. Perhaps, by distributing benefits on racial lines, such policies impugn people's dignity and autonomy, promote racial depreciation and essentialism, and drive racially divisive politics. Or, by recognizing and accounting for individuals' otherwise overlooked ability, they might promote dignity and autonomy, reduce racial essentialism and depreciation, and deprive political backlash of its moral force. For the moment, what can be confidently stated is that these questions await answers.

From the perspective of schools seeking to implement affirmative meritocracy, our legal analysis suggests the maxim “set in order thy house.” Schools that reduce threat will unleash their students’ latent ability and thereby preempt potential legal liability. They can do so at modest cost and without employing legally risky racial classifications.

By contrast, schools will find it difficult to cheaply, effectively, and legally correct for effects of stereotype threat in prior education environments on measures used for selection. The most accurate and affordable approach—mechanical, race-based score recalibration—raises the hardest legal questions. Individualized consideration carries reduced risks and reduced benefits. Given that schools will be held responsible if they incorrectly read the Court’s tea leaves concerning which approach best satisfies narrow tailoring, some may abandon the field. The result: non-Asian minority applicants would be declined admission in favor of white and Asian applicants whom they could predictably outperform.

In any case, affirmative meritocracy promises much. Policies that reduce stereotype threat have the potential to eliminate an important source of U.S. racial disparities in education in legal ways consistent with reigning ideals of antidiscrimination law.