



Case Western Reserve University School of Law Scholarly Commons

Faculty Publications

2020

Holistic Review in Race-Conscious University Admissions

Hal Arkes

George W. Dent Jr.

Case Western Reserve University, george.dent@case.edu

Follow this and additional works at: https://scholarlycommons.law.case.edu/faculty_publications

 Part of the [Constitutional Law Commons](#)

Repository Citation

Arkes, Hal and Dent, George W. Jr., "Holistic Review in Race-Conscious University Admissions" (2020).
Faculty Publications. 2039.

https://scholarlycommons.law.case.edu/faculty_publications/2039

This Article is brought to you for free and open access by Case Western Reserve University School of Law Scholarly Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Case Western Reserve University School of Law Scholarly Commons.

HOLISTIC REVIEW IN RACE-CONSCIOUS UNIVERSITY ADMISSIONS

Introduction

The Supreme Court has held that race may be considered as “a factor of a factor of a factor” within a “holistic” program of university admissions if the university can satisfy a heavy burden of proving that the program is “narrowly tailored” to achieve the educational benefits of diversity.¹ The Court has listed the desired benefits of racial diversity, but it has not discussed what evidence a university needs to prove that its program is “narrowly tailored” to achieve those benefits.

This article addresses that issue. The field of psychology offers abundant research about the process of judgment and decision-making (“JDM”) and testing the validity of any particular program of decision-making. This knowledge offers valuable insights about how a university can design and, later, measure the validity of its admissions process so that it is “narrowly tailored” to achieve the educational benefits of racial diversity approved by the Supreme Court.

Part I describes the standard that the Court has laid down for race-conscious admissions programs. Part II reviews the JDM literature on holistic and disaggregated decision-making. Part III discusses how the conclusions of this literature can guide the courts in applying the standard of review laid down by the Supreme Court.

I. The Supreme Court’s Standard for Race-Conscious University Admissions

In *Regents of the University of California v. Bakke* the Supreme Court held that “attainment of a diverse student body ... is a constitutionally permissible goal for an institution of higher education.”² Subsequent cases confirmed the constitutional validity of this goal.³

In *Fisher v. University of Texas*⁴ Abigail Fisher, a student rejected for admission to the University of Texas – Austin (“UT”), challenged the school’s admissions system as racially discriminatory in violation of the Equal Protection Clause of the 14th Amendment.⁵ Although

¹ *Fisher v. Univ. of Texas*, 136 S.Ct 2198, 2207 (2016) (*Fisher II*).

² *Regents of Univ. of Calif. v. Bakke*, 438 U.S. 265, 311 (1978).

³ *Grutter v. Bollinger*, 539 U.S. 306, 328 (2003) (“Today, we hold that the Law School has a compelling interest in attaining a diverse student body.”).

⁴ *Fisher II*; *Fisher v. Univ. of Texas*, 133 S.Ct. 2411 (2013) (*Fisher I*).

⁵ U.S. CONST. amend. XIV.

UT admitted that it considered race in its admissions decisions, it claimed that it did so as just one of many factors in a holistic review carefully designed to obtain the educational benefits of diversity as sanctioned by *Bakke*.⁶

In its first encounter with the case (“*Fisher I*”) the Court reiterated the validity of diversity as an educational goal: “The attainment of a diverse student body . . . serves values beyond race alone, including enhanced classroom dialogue and the lessening of racial isolation and stereotypes.”⁷ In *Grutter v. Bollinger* the Court had stated that “a permissible goal . . . require[s] only a good-faith effort . . . to come within a range demarcated by the goal itself.”⁸ However, in *Fisher I* the Court found that

the District Court and the Court of Appeals had “confined the strict scrutiny inquiry in too narrow a way by deferring to the University’s good faith in its use of racial classifications.” . . . The Court remanded the case, with instructions to evaluate the record under the correct standard and to determine whether the University had made “a showing that its plan is narrowly tailored to achieve” the educational benefits that flow from diversity. . . . On remand, the Court of Appeals determined that the program conformed with the strict scrutiny mandated by *Fisher I*.⁹

Reviewing this decision in *Fisher II*, the Supreme Court affirmed the decision of the Fifth Circuit. In so doing, it said: “Once a university gives ‘a reasoned, principled explanation’ for its decision, deference must be given ‘to the University’s conclusion, based on its experience and expertise, that a diverse student body would serve its educational goals.’”¹⁰ The Court reiterated its statement in *Grutter* that enrolling a diverse student body “promotes cross-racial understanding, helps to break down racial stereotypes, and enables students to better understand persons of different races” and that “student body diversity promotes learning outcomes, and better prepares students for an increasingly diverse workforce and society.”¹¹

However: “Strict scrutiny requires the university to demonstrate with clarity that its ‘purpose or interest is both constitutionally permissible and substantial, and that its use of the classification is necessary . . . to the accomplishment of its purpose.’”¹² Moreover: “*Fisher I* clarified that no deference is owed when determining whether the use of race is narrowly tailored

⁶ *Fisher II*, 136 S.Ct at 2207.

⁷ *Fisher I*, 133 S.Ct. at 2417-18.

⁸ *Grutter v. Bollinger*, 539 U.S. 306, 335 (2003) (quoting *Sheet Metal Workers v. EEOC*, 478 U.S. 421, 495 (1986)).

⁹ *Fisher II*, 136 S.Ct. at 2208 (internal citations omitted).

¹⁰ *Fisher II*, 136 S.Ct. at 2208. *See also* *Grutter*, 539 U.S. at 328 (“The Law School’s educational judgment that such diversity is essential to its educational mission is one to which we defer.”).

¹¹ *Id.* at 2210, quoting *Grutter*, 539 U.S. at 330.

¹² *Id.* at 2208, quoting *Fisher I*, 133 S.Ct. at 2418. *See also* *Grutter*, 539 U.S. at 326-27 (“all governmental uses of race are subject to strict scrutiny”).

to achieve the university’s permissible goals.”¹³ And, “asserting an interest in the educational benefits of diversity writ large is insufficient. A university’s goals cannot be elusory or amorphous—they must be sufficiently measurable to permit judicial scrutiny of the policies adopted to reach them.”¹⁴ “The purpose of the narrow tailoring requirement is to ensure that ‘the means chosen “fit” . . . th[e] compelling goal so closely that there is little or no possibility that the motive for the classification was illegitimate racial prejudice or stereotype.’”¹⁵

This is consistent with the Court’s statement in *Fisher I* that “‘the mere recitation of a “benign” or legitimate purpose for a racial classification is entitled to little or no weight.’ . . . Strict scrutiny does not permit a court to accept a school’s assertion that its admissions process uses race in a permissible way without a court giving close analysis to the evidence of how the process works in practice.”¹⁶

In an effort to meet this standard, UT utilized a “holistic” admissions review process comprising several factors, including “personal achievement . . . test scores, or other unique skills” and “differences in life experiences.”¹⁷ Among these different experiences was “the experience of being a minority in a majority-white or majority-minority school and succeeding in that environment [which] offers a rich pool of potential UT Austin students with demonstrated qualities of leadership and sense of self.”¹⁸ In *Fisher II* the Court stated that

the University identifies the educational values it seeks to realize through its admissions process: the destruction of stereotypes, the “ ‘promot[ion of] cross-racial understanding,’ ” the preparation of a student body “ ‘for an increasingly diverse workforce and society,’ ” and the “ ‘cultivat[ion of] a set of leaders with legitimacy in the eyes of the citizenry.’ ” . . . Later in the proposal, the University explains that it strives to provide an “academic environment” that offers a “robust exchange of ideas, exposure to differing cultures, preparation for the challenges of an increasingly diverse workforce, and acquisition of competencies required of future leaders.” . . . All of these objectives, as a general matter, mirror the “compelling interest” this Court has approved in its prior cases.¹⁹

¹³ *Id.* at 2208, citing *Fisher I* at 2419-20. *See also Grutter*, 539 U.S. at 326 (“such classifications are constitutional only if they are narrowly tailored to further compelling governmental interests”).

¹⁴ *Id.* at 2211.

¹⁵ *Grutter*, 539 U.S. at 333 (quoting *Richmond v. J.A. Croson Co.*, 488 U.S. 469, 493 (1989) (plurality opinion)).

¹⁶ *Fisher I* 133 S.Ct. at 2421.

¹⁷ *Fisher v. Univ. of Texas*, 758 F.3d 633, 653 (2014).

¹⁸ *Id.*

¹⁹ *Fisher II*, 136 S.Ct. at 2211 (internal citations omitted).

The Supreme Court affirmed the district court's finding that, in UT's admissions decisions, "race is but a 'factor of a factor of a factor.'"²⁰ Further, the Court found that the plaintiff had failed to identify any race-neutral admissions process that could fulfill UT's diversity goals.²¹ It therefore upheld the constitutionality of UT's admissions program.

The scope of the Court's review of that program was circumscribed, however, by the plaintiff's limited challenge to it. The dissenting opinion by Justice Alito noted several problems with the studies presented by UT to document the educational benefits of its holistic review process.²² However, the majority noted that Fisher had not challenged the adequacy of these studies, and it declined to consider the evidence cited by Justice Alito's dissenting opinion because, *inter alia*, UT had had no opportunity to respond to that evidence.²³

Among the issues that the plaintiff's limited challenge avoided was the meaning of the concept of a "critical mass" of minority students which UT claimed that it needed. The majority stated that plaintiff argued that "the University must set forth more precisely the level of minority enrollment that would constitute a 'critical mass.'"²⁴ The majority viewed this as a demand for a precise number, which it considered improper because "the University is prohibited from seeking a particular number or quota of minority students."²⁵

Likewise, the Court never discussed just how much weight UT gave to race. In *Grutter* the Court said that race may not be used "in a way that makes an applicant's race or ethnicity the defining feature of his or her application."²⁶ The majority in *Fisher II* found that in UT's process "race is but a 'factor of a factor of a factor,'"²⁷ but it said nothing about how much weight that factor was actually given. Evidently plaintiff did not raise that question. *A fortiori*, the Court did not discuss how much weight *may* be given to race.

Further, the Court did not inquire whether UT used non-racial admissions criteria in a race-neutral way. In *Grutter* the Court said that the University of Michigan Law School gave consideration for, *inter alia*, "admittees who have lived or traveled widely abroad, are fluent in

²⁰ *Fisher II*, 136 S.Ct at 2208, quoting the district court opinion, 645 F. Supp. 2d 587, 608 (W.D. Tex. 2009).

²¹ 136 S.Ct. at 2211-12.

²² 136 S.Ct. at 2215-16, 2219-20, 2226-29, 2239-40 (Alito, J., dissenting).

²³ *Fisher II*, 136 S.Ct. at 2211-12. It seems that the plaintiffs in *Grutter* also failed to raise the issue of the empirical validity of the school's use of race in admissions in order to attain educational benefits. The majority opinion in *Grutter* discussed the design of the school's admissions reviews. 539 U.S. at 334-41. However, the Court mentioned no studies to determine whether the school's program had actually achieved the intended benefits of diversity.

²⁴ *Fisher II*, 136 S.Ct. at 2210.

²⁵ *Id.*

²⁶ *Grutter*, 539 U.S. at 337.

²⁷ *Fisher II*, 136 S.Ct at 2207, quoting the district court opinion, 645 F. Supp. 2d 587, 608 (W.D. Tex. 2009).

several languages, have overcome personal adversity and family hardship, have exceptional records of extensive community service, and have had successful careers in other fields.”²⁸ UT did not show whether these factors were applied in a race-neutral way; how much weight these factors received; or whether in practice educational benefits were actually realized by inclusion of these factors.

The majority tacitly acknowledged the limited scope of its holding in *Fisher II*. While it sustained UT’s program, it imposed on the university a continuing duty, *inter alia*, “to identify the effects, both positive and negative, of the affirmative-action measures it deems necessary. . . . It is the University’s ongoing obligation to engage in constant deliberation and continued reflection regarding its admissions policies.”²⁹

Fisher II thus leaves open how a court should proceed if a plaintiff questions the educational benefits of a particular university’s race-conscious admissions program. *Fisher II* and its predecessors make clear that the program is subject to “strict scrutiny” and that “no deference is owed when determining whether the use of race is narrowly tailored to achieve the university’s permissible goals.”³⁰ Further, those goals “must be sufficiently measurable to permit judicial scrutiny of the policies adopted to reach them.”³¹ What evidence must a university produce to satisfy this standard? Part II looks to the research on judgment and decision-making for insights about how a selection process can best be designed and later checked for validity. Part III applies this research to university admissions programs.

II. University Admission Procedures in Light of the Judgment and Decision-Making Literature

A. Holistic ratings

The judgment and decision-making [“JDM”] literature on evaluation makes a distinction between holistic ratings and disaggregated ratings. As an example, the Olympic ice skating competition was once judged using two criteria: presentation and technical merit, each criteria being judged on a 1 to 6 scale. Because each criterion was scored separately, this would be termed a “disaggregated” rating scheme. In such schemes the disaggregated ratings are eventually combined in some mathematical manner to yield a final score. For example, each of the two disaggregated ratings could first be multiplied by a weight to indicate the importance the Olympic committee thought should be attached to each criterion, and then those two resulting products could be added together. Had the judges instead provided only a single score that reflected their attempt to take into account both criteria in some way, this would have been a “holistic” rating scheme.

²⁸ *Grutter*, 539 U.S. at 338.

²⁹ *Id.* at 2215.

³⁰ *Id.* at 2208.

³¹ *Id.* at 2211.

The University of Texas at Austin has used a “Personal Achievement Score” (PAS) in its evaluation of applicants. This score is graded on a 1 to 6 scale, and is “. . . determined by a separate reader, who (1) rereads the applicant’s required essays, (2) reviews any supplemental information the applicant submits (letters of recommendation, resumes, an additional optional essay, writing samples, artwork, etc.), and (3) evaluates the applicant’s potential contributions to the University’s student body based on the applicant’s leadership experience, extracurricular activities, awards/honors, community service, and other ‘special circumstances.’ ‘Special circumstances’ include the socioeconomic status of the applicant’s family, the socioeconomic status of the applicant’s school, the applicant’s family responsibilities, whether the applicant lives in a single-parent home, the applicant’s SAT score in relation to the average SAT score at the applicant’s school, the language spoken at the applicant’s home, and, finally, the applicant’s race.”³² All of these factors are potential inputs into the 1-6 PAS score. Thus the PAS certainly qualifies as a holistic rating. “Holistic ratings” in the research literature does not mean “considering relevant factors other than academic credentials.” The research literature contains no objections whatsoever to that strategy. However the literature is very skeptical about the wisdom of using holistic ratings that amalgamate unmeasured components into a single score. The following are the major conclusions derived from the research literature concerning holistic versus disaggregated ratings.

Conclusion #1: Disaggregated ratings often make more accurate predictions than do holistic ratings.

A very substantial amount of JDM research has compelled the conclusion that the use of disaggregated ratings often makes significantly more accurate predictions than do holistic ratings.³³ Grove et al. performed a meta-analysis, which is a technique that combines the results of a multitude of research papers pertaining to the same topic. The researchers combined 136 studies in their meta-analysis in order to compare the ability of holistic and disaggregated ratings to predict outcomes in an extremely wide variety of fields including medicine, student admission decisions, sales, and many other fields. The results were as follows: In approximately 46% of the studies examined, depending on the specific analysis, the process of arithmetically combining disaggregated ratings outperformed intuitive holistic prediction. For 48% of the studies the two techniques yielded the same performance. For 6% of the studies the intuitive holistic procedure was superior, although the majority of these were superior by only a very small amount.

One of the first demonstrations of the superiority of disaggregated ratings over holistic ratings was done by Dawes.³⁴ As a member of the University of Oregon Department of Psychology, Dawes had access to the holistic ratings the clinical psychology faculty had given to applicants to that graduate program. He also had access to the faculty’s rating of each accepted

³² Fisher v. Univ. of Texas (*Fisher II*), 136 S. Ct. 2206 (2016).

³³ William M. Grove et al., *Clinical Versus Mechanical Prediction: A Meta-analysis*, 12 PSYCHOL. ASSESS. 19 (2000).

³⁴ Robyn M. Dawes, *A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making*, 26 AM. PSYCH. 180 (1971).

student's progress after the first year of graduate school. Note that the admissions committee contained only Ph.D. professional clinical psychologists, who are professionally trained to evaluate personal factors—more trained than are members of any college admissions committee. Nevertheless, the correlation between the rating given to each accepted student's application and the evaluation of the student after a year in the program was only .19. This means that the clinicians' rating of the admission materials accounted for only about 3.6% of the variation among the students' first-year graduate school performance. However, a simple amalgamation of each student's GPA, Graduate Record Examination score, and the quality of the undergraduate institution attended was correlated .38 with each student's first-year graduate school performance. Thus when these three disaggregated scores were combined, they *quadrupled* the amount of variance accounted for compared to the clinicians' holistic ratings of the applications.³⁵ This means that the amalgamation of the three disaggregated scores was a drastically better predictor of the students' first year evaluation than were the clinicians' ratings of the admissions materials. Several studies examining personal evaluations and predictions, including many pertaining to the accuracy of predictions made in legal testimony, have come to the same conclusion.³⁶ For years scholars have commented on the stark contrast between the Supreme Court's insistence on holistic evaluations and the research literature's denigration of such ratings.³⁷

The first and most important conclusion supported by the JDM literature is that disaggregated ratings actuarially amalgamated by a simple mathematical formula often are superior to holistic ratings, even when the holistic ratings are done by experts. We wish to emphasize that such analyses are easy to do. Undergraduate research courses employ “canned” programs that perform such analyses.

Conclusion #2: Decision makers are insufficiently aware of the bases for their decisions.

In the JDM literature “self-insight” is a term that refers to the extent to which a decision maker is able to articulate accurately the bases for his or her decision. The research literature suggests that self-insight is very poor. We emphasize that we do not imply that when a decision

³⁵ By squaring the correlation one obtains the amount of variance accounted for: $4(.19)^2 = (.38)^2$.

³⁶ Robyn M. Dawes et al., *Clinical Versus Actuarial Judgment*, 243 *SCIENCE* 1668 (1989).

³⁷ See Margaret E. Brooks et al., *Distinction Bias in Applicant Reactions to Using Diversity Information in Selection*, 17 *INT'L J. SELECT. & ASSESS.* 377, 378 (2009): “In the most Supreme Court case dealing directly with this issue, the Court mandated a more holistic approach to affirmative action—despite evidence that mechanical methods of combining data consistently outperform holistic approaches.” See also Scott Highhouse & John A. Kostech, *Holistic Assessment for Selection and Placement*, in K. F. Geisinger (Ed.), *APA HANDBOOK OF TESTING AND ASSESSMENT IN PSYCHOLOGY: VOL I. TEST THEORY AND TESTING AND ASSESSMENT IN INDUSTRIAL AND ORGANIZATIONAL PSYCHOLOGY* 565 (2013).

maker is inaccurate in describing his or her decision-making policy, he or she is necessarily being duplicitous. The literature suggests a lack of awareness, not a lack of honesty.³⁸

Examples abound. Elstein and colleagues³⁹ asked a group of 50 physicians to participate in research designed to ascertain when hormonal replacement therapy should be used for post-menopausal women. The group of 50 included 25 gynecologists and 25 family physicians. Each physician rated their own likelihood of prescribing estrogen to women in each of 12 case descriptions. Each physician also provided their estimates of the probability of each of the possible outcomes, the utility of each of those outcomes, and the importance of each of the outcome categories. From each physician's ratings of these factors it was easy to ascertain what estrogen replacement recommendation should be compelled for each case by each physician. The results were that the recommendations based on the physician's own ratings of the various factors were poorly related to the recommendation actually rendered by each physician for each of the 12 cases! In other words, the final decision was not well described by each physician's own stated policy.

The analogous finding has also been found in legal contexts. For example, Dhami and Ayton⁴⁰ examined the bail versus jail decisions of 81 magistrates in the United Kingdom. Each magistrate stated what factors he or she considered in making each decision and how much weight was placed on each factor. From a simple regression analysis the authors were able to ascertain which factors actually influenced each magistrate's decision and how much weight was placed on each factor. The results were that the bases of the actual decisions corresponded poorly to both the stated policy and to "the ideal practice as defined by the due process model of justice."⁴¹ The magistrates could not accurately articulate the true bases of their own decisions. The analogous result has also been found when abstract stimuli were used and where there would therefore be no reason to be dishonest or motivation to conform to any particular policy.⁴²

Conclusion #3: Disaggregating a holistic rating allows one to discern the policy underlying the holistic rating; thus disaggregated ratings promote transparency

Given that it is so difficult to articulate accurately the bases of one's evaluations and decisions, disaggregating a holistic rating into its components can perform the very valuable

³⁸ Richard E. Nisbett & Timothy D. Wilson, *Telling More Than We Can Know: Verbal Reports on Mental Processes*, 84 PSYCH. REV. 231 (1977).

³⁹ Arthur S. Elstein et al., *Comparison of Physicians' Decisions Regarding Estrogen Replacement Therapy for Menopausal Women and Decisions Derived from a Decision Analytic Model*, 80 AM. J. MED. 246 (1986).

⁴⁰ Mandeep K. Dhami & Peter Ayton, *Bailing and Jailing the Fast and Frugal Way*, 14 J. BEHAV. DECISION MAKING 141 (2001).

⁴¹ *Id.* at 141.

⁴² Paul Slovic & Sarah Lichtenstein, *Comparison of Bayesian and regression approaches to the study of information processing in judgment*, 6 ORGAN. BEHAV. HUMAN PERF. 649 (1971).

function of revealing which components—if any—are responsible for the holistic rating. The following example illustrates this point in an educational setting.

The National Science Foundation (NSF) provides funding for doctoral students to complete their dissertations.⁴³ Students submit proposals, and within each discipline a panel of experts is convened to evaluate each proposal and decide which ones merit funding. Such financial support is often essential for graduate students who have to do fieldwork in a foreign country, for example.

The General Accounting Office (GAO) had criticized NSF and two other funding agencies for allowing informal and unwritten criteria to pollute panels’ funding recommendations.⁴⁴ Dissertation proposals were *supposed* to be evaluated according to the following four criteria: methodology, quality of the student’s training, utility of the research, and the theoretical basis of the research.⁴⁵ Panelists had to provide only one holistic rating rather than provide a rating indicating how well each proposal satisfied each of the four criteria. I (HRA) asked the relevant program officer to ask the panelists to do both types of ratings: provide a holistic rating for each proposal and provide disaggregated ratings for each criterion. Using a simple regression analysis, I merely regressed the ratings of each panelist on the individual criteria on that panelist’s overall rating. Table 1 contains the results of this analysis. The “Multiple R” column indicates that the holistic rating of three of the four panelists

Table 1. Use of four criteria by members of a dissertation award panel at the National Science Foundation

Panelist	Criterion								Multiple R
	Methodology		Student's training		Theory		Utility		
	Beta	Significance	Beta	Significance	Beta	Significance	Beta	Significance	
1	.806	<.001	.108	<.01	.026	n.s.	.111	n.s.	.974
2	.332	<.01	.103	n.s.	.130	n.s.	.201	n.s.	.537
3	.381	<.001	.280	<.01	.120	n.s.	.244	<.05	.895
4	.445	<.001	.215	<.001	.076	n.s.	.334	<.001	.968

Note. The table shows results of an analysis in which the ratings of each panelist on the individual criteria were regressed on that panelist’s overall rating.

was overwhelmingly determined by the ratings given to the four criteria, although not all four criteria’s ratings contributed significantly to the holistic one. In particular, no panelist’s rating of a proposal’s theory contributed significantly to their overall holistic rating. Note that Panelist #2’s ratings of the four criteria were far less deterministic of the holistic rating than was the case for the other panelists. The program officer had been suspicious of this panelist’s rating performance, and this analysis confirmed his suspicions. There were two possible reasons why Panelist #2’s criterion ratings and holistic rating were not closely related. First, Panelist #2 might

⁴³ Special Programs for Graduate Students, at https://www.nsf.gov/funding/education.jsp?fund_type=2.

⁴⁴ United States General Accounting Office, *Peer Review: Reforms Needed to Ensure Fairness in Federal Agency Grant Selection*, GAO/PEMD-94-1 at 83.

⁴⁵ Reprinted from Hal R. Arkes, *The Nonuse of Psychological Research at Two Federal Agencies*, 14 PSYCH. SCI. 1 (2003).

have been trying to use the four legitimate criteria but did so inconsistently. Second, this panelist may have been basing her holistic rating on illegitimate criteria, a prohibited behavior that had drawn the disapproval of the GAO. By using disaggregated ratings, the program officer was able to discern whether and to what extent every panelist was using each of the four criteria approved by the National Science Board. This is another enormous advantage of disaggregated over holistic ratings.

Compare the transparency of this disaggregated analysis with the procedure of The University of Texas at Austin (UTA), the defendant in *Fisher*. UTA uses a “Personal Achievement Score” (PAS) to help determine who will be admitted to the University. As mentioned earlier, this score includes demonstrated leadership qualities, extracurricular activities, honors and awards, work experience, community service and a variety of special circumstances. The special circumstances in turn include the socioeconomic status of the applicant’s family (parents’ education and combined income); whether the applicant is from a single-parent home; whether a language other than English is primarily spoken in the applicant’s home; special family responsibilities; the socioeconomic status of the applicant’s high school; the average SAT/ACT score of students attending the applicant’s high school in relation to the applicant’s own score; and beginning with applications for the 2005 entering class, the applicant’s race or ethnicity.⁴⁶

If a reader of an applicant’s file gives a PAS rating of 4, for example, it is not possible to know what might be the basis for that rating. Was the work experience particularly influential in deciding upon that rating? Perhaps it was the applicant’s race. Perhaps the reader detected a propensity for leadership. It is impossible to know. However, if the reader were forced to provide disaggregated ratings for each component of the PAS, it would be readily apparent what the bases of the holistic rating would be. Disaggregated ratings strongly promote transparency. Holistic ratings are completely opaque with regard to the influence of any component upon the holistic rating.

B. Possible objections to the use of disaggregated ratings

Given the advantages of disaggregated ratings, or equivalently, the disadvantages of holistic ratings, what could be possible objections to the use of the former over the latter?

Possible Objection #1: Numbers are dehumanizing.

This has been the traditional objection to the use of numerical ratings on any aspect of a person’s characteristics or behavior.⁴⁷ However the University of Texas, for example, *already* uses numerical ratings, such as the 1-6 PAS score, among others. It is true that disaggregating a holistic rating into its components requires more numbers than does a single holistic rating. Unless the number of digits used is linearly related to the magnitude of dehumanization,

⁴⁶ See *supra* note 32 and accompanying text.

⁴⁷ William M. Grove & Paul E. Meehl, *Replies to Commonly Heard Objections to Actuarial Judgment*, 2 PSY. PUB. POLICY & L. 293 (1996).

disaggregation should not result in any more dehumanization than does the current usage of ratings by universities. It should be acknowledged that making better decisions about applicants is more humanizing than making inaccurate decisions about applicants, and the research evidence is quite clear that disaggregated ratings generally foster more accurate decisions.

Possible Objection #2: One cannot put numbers on one's private thoughts.

It strikes some people as inappropriate to assign ratings to components of a holistic rating. Some people think that the holistic rating comes about through a cognitive process that is inaccessible to conscious consideration and measurement. However, one's unconscious consideration to create a holistic rating and one's conscious assignment of a number to each component are functionally identical. The only difference is that the former is done implicitly and the latter is done explicitly. When one mentally creates a single holistic rating, one is tacitly considering each component to some unknown degree. The disaggregated technique simply requires that one express what one normally does implicitly. This expression makes the process transparent and thereby accessible to others and even to oneself.

Possible Objection #3: Disaggregated ratings don't take into account statistical interactions.

This is an esoteric statistical objection. Suppose that in the eyes of an admissions officer an applicant's being in many extracurricular activities would be a "plus" if the applicant did not have to earn income to care for other family members, but would be a "minus" if the applicant did have to earn income to care for other family members. This would comprise a statistical interaction between these two factors of the PAS; the presence (or absence) of a factor (extracurricular participation) would mean something different depending on the level of another factor (care of family members). Although a multiple regression analysis can detect such interactive relations between multiple variables, critics might argue that interactions between more than two variables would become so statistically complicated that it would be better to rely on each admission officer's expert holistic consideration of such complexities. However research has shown that human decision makers are unable to discern or use such complicated relations to any substantial extent when evaluating stimuli.⁴⁸ Therefore the basis for this objection has been undermined by empirical research.

Possible Objection #4: People prefer holistic ratings; therefore they should be used.

There is no doubt that people prefer holistic ratings over disaggregated ones.⁴⁹ However it is irrational to prefer a technique that makes less accurate predictions over one that makes more accurate predictions.

⁴⁸ Berndt Brehmer & A. Brehmer, *What Have We Learned About Human Judgment From Thirty Years of Policy Capturing?* in B. Brehmer and C.R. Joyce (Eds.), *HUMAN JUDGMENT: THE SOCIAL JUDGMENT THEORY VIEW* 75 (1988).

⁴⁹ Hal R. Arkes et al., *Assessing the Merits and Faults of Holistic and Disaggregated Judgments*, 23 *J. BEHAV. DECISION MAKING* 250 (2010). See also Joseph Eastwood et al., *What People Want From Their*

At least two reasons exist why people prefer the less accurate technique. First, people think that they have more control over holistic ratings. With disaggregated ratings, the final amalgamated rating is compelled by the components' ratings. Occasionally people don't like the result of the amalgamation. Raters may be unpleasantly surprised when their own ratings of the components yield an unwanted or unexpected result. Second, holistic ratings allow people to effectively disguise the bases for their decisions. If Panelist #2 in the NSF example above were basing her rating on some illegitimate factor, the holistic rating would be an ideal way to hide this corrupt behavior. On the other hand, with disaggregated ratings Panelist #2 could be asked, for example, why her rating of the dissertation methodology was so low. With disaggregated ratings the decision to downgrade a dissertation proposal would have to be manifested in a low rating given to at least one of the components, and a justification for this low rating could then be requested.

C. Self-Report Measures

Many of the studies purportedly supportive of diversity have used self-report measures. The serious problem with such research can be exemplified by a pair of studies by Conway and Ross.⁵⁰ The authors assessed the impact of a study skills course in two different experiments. The course was totally worthless with regard to improving any measure of academic performance. However, students in both experiments (a) misremembered their starting skills as worse than they really were and (b) judged their study skills to have improved significantly thanks to the course. The students assumed that by taking a study skills course their study skills had improved.

Now consider the amicus brief by Professor Gurin filed in the two Michigan cases.⁵¹ In the Cooperative Institutional Research Program research study discussed in the amicus brief, diversity was defined by a student's enrollment in an ethnic studies course in college. These students were asked to self-report changes in such factors as general knowledge and ability to think critically. As veteran university professors, we suggest that students are not particularly accurate judges of whether they are generally knowledgeable and able to think critically. (If they were, there would be no need for any tests.) As was the case for the participants in the study skills course examined by Conway and Ross, the participants in the ethnic studies course were not in an unbiased position to assess their growth in various areas. Naturally the students assumed that they had improved. In general, unverified self-report measures of improvement are

Professionals: Attitudes Toward Decision-making Strategies, 25 J. BEHAV. DECISION MAKING 458 (2012).

⁵⁰ Michael Conway & Michael Ross, *Getting What You Want By Revising What You Had*, 47 J. PERSONAL & SOC. PSYCH. 738 (1984).

⁵¹ *Gratz. v. Bollinger*, 539 U.S. 244 (2003); *Grutter v. Bollinger*, 539 U.S. 306 (2003).

fraught with difficulties, especially when improvement represents a socially desirable response that the respondent knows will be well received.⁵²

There is an irony in relying on self-reports to support the diminution of prejudice purportedly caused by diversity. Many researchers in the domain of prejudice insist that relying on self-report measures of one's own prejudice are poor indicators of true bigotry, because people are either unwilling to disclose or unaware of the presence of their own prejudice.⁵³ However proponents of diversity frequently rely on self-report measures to support the purported benefits of diversity.

III. Do Current Policies Satisfy the Law?

To reiterate, the Supreme Court says that race discrimination in school admissions is subject to strict scrutiny. "Strict scrutiny requires the university to demonstrate with clarity that its 'purpose or interest is both constitutionally permissible and substantial, and that its use of the classification is necessary . . . to the accomplishment of its purpose.'"⁵⁴ Further, the use of race must be "narrowly tailored to achieve the university's permissible goals."⁵⁵ "The purpose of the narrow tailoring requirement is to ensure that 'the means chosen "fit" . . . th[e] compelling goal so closely that there is little or no possibility that the motive for the classification was illegitimate racial prejudice or stereotype.'"⁵⁶

The majority in *Fisher II* did not consider whether UT satisfied this standard because plaintiff did not raise the issue.⁵⁷ The judgment and decision-making ("JDM") literature shows

⁵² See Douglas P. Crowne and David Marlowe, *A New Scale of Social Desirability Independent of Psychopathology*, 24 J. CONSULT PSYCH. 349 (1960); Thomas E. Wood & Malcolm J. Sherman, *Supplement to Race and Higher Education: Why Justice Powell's Diversity Rationale for Racial Preferences in Higher Education Must Be Rejected*, National Association of Scholars, <https://files.eric.ed.gov/fulltext/ED455755.pdf>, at p. 4: "Students who take ethnic studies courses or attend a racial workshop may be more likely to know what the approved attitudes or behaviors are, and thus to report them."

⁵³ Brian A. Nosek, et al., *Math = Me, Me = Female, Therefore Math Is Not Equal To Me*, 83 J. PERSONAL & SOC. PSYCH. 44, 55 (2002): "A stereotype may be maintained outside conscious awareness although it is neither wanted nor endorsed consciously, yet still influence both consciously and unconsciously held attitudes."

⁵⁴ *Fisher II* at 2208, quoting *Fisher I* at 133 S.Ct. at 2418.

⁵⁵ *Id.* at 2203, citing *Fisher I* at 2419-20. See also *Grutter*, 539 U.S. at 326 ("such classifications are constitutional only if they are narrowly tailored to further compelling governmental interests").

⁵⁶ *Grutter*, 539 U.S. at 333 (quoting *Richmond v. J.A. Croson Co.*, 488 U.S. 469, 493 (1989) (plurality opinion)).

⁵⁷ *Fisher II*, 136 S.Ct. at 2211-12.

that methods like those used in *Fisher* do not meet this standard and therefore do not satisfy strict scrutiny.

A. Holistic Ratings Do Not “Demonstrate with Clarity” but Instead Obscure Underlying Policies

Because UT Austin uses holistic rather than disaggregated ratings, the actual admissions policy is largely opaque. Thus detailed judicial scrutiny of the policy is impossible. We are not the first to note this deficiency. Zell bluntly stated that “. . . the sole guiding principle behind the holistic-admissions system is that this system allows colleges to accept—and to reject—whomever they desire and then to hide the reasons for their decision.”⁵⁸ Justice Ginsburg hinted at this feature of holistic ratings: “As for holistic review, if universities cannot explicitly include race as a factor, many may ‘resort to camouflage’ to ‘maintain their minority enrollment.’”⁵⁹ Holistic ratings are ideally suited for “camouflage.” UCLA law professor Richard Sander explained how many law school faculty wanted to exploit the opacity of holistic ratings after the citizens of California passed Proposition 209 banning racial preferences: “Some faculty observed that admissions decision in many graduate departments rested on so many subjective criteria that it would be easy to make the continued consideration of race invisible to outsiders.”⁶⁰ Holistic evaluations eliminate any easy way to detect such defiance of Proposition 209 or any other law or court ruling.

Zell noted that when Jewish applicants started outscoring “blue blood” applicants in the 1920s, the elite Eastern universities began to use non-academic criteria such as “character” to screen out the former in favor of the latter, a procedure currently alleged to be used against Asian students.⁶¹ Harvard President Lowell advised against applying “character” to both Jewish and Gentile applicants.⁶² “Both to employ a double standard and then to hide its use, a university could find no more effective tool than the holistic admissions system.”⁶³ With disaggregated ratings it would have been extremely easy, for example, to ascertain if Gentiles were disproportionately given extra points for “character.” Today it would similarly be extremely easy to ascertain if students from one demographic group are disproportionately given extra

⁵⁸ Jonathan R. Zell, *It’s Not About Race: The True Purpose of the University of Texas’ Holistic-Admissions System Is to Give Preferences to Well-Connected White Applicants, Not To Disadvantaged Minorities*, 24, U. MIAMI BUS. L. REV. 35, 68 (2016).

⁵⁹ *Fisher I*, 133 S. Ct. at 2433 n.2 (Ginsburg, J., dissenting) (quoting *Gratz v. Bollinger*, 539 U. S. 244, 304 (Ginsburg, J., dissenting)). Justice Souter also suggested that universities may try to obfuscate their illegitimate use of race. *Gratz v. Bollinger*, 539 U.S. 244, 298 (2003) (Souter, J., dissenting).

⁶⁰ Heather MacDonald, *California Passed an Anti-Affirmative Action Law, and Colleges Ignored It*, <https://nypost.com/2018/09/01/california-passed-an-anti-affirmative-action-law-and-colleges-ignored-it/> downloaded 3/2/19.

⁶¹ Zell, *supra* note 58, at 57.

⁶² *Id.* at 69.

⁶³ *Id.* at 76

points on non-academic characteristics in order to inflate their Personal Achievement Scores (“PAS”). UT admitted that if it had relied on standardized test scores alone, almost all of the applicants admitted would have been White.⁶⁴ This implies that minorities were rated higher than Whites on the non-academic factors that were the inputs into the PAS. With disaggregated ratings it would be easy to ascertain if, compared to Whites, minorities were deemed to have manifested much more leadership, for example. In a current lawsuit Asian students assert that non-academic factors are being used to diminish their chances of being admitted to Harvard.⁶⁵ With disaggregated ratings, the impact of each non-academic factor can be unambiguously assessed.

In addition to providing greater transparency, disaggregated ratings offer better institutional control over the ratings process. With holistic ratings, the admissions administration cannot tell with any precision how much weight each rater gives to each factor; the raters, who are students or low-level employees, have wide discretion. However, universities claim that they have a “compelling interest” in attaining the educational benefits of diversity, as indeed they must in order to satisfy the constitutional requirements for engaging in race-conscious admissions.⁶⁶ If affirmative action is so important, administrators should exercise tight control over how it is implemented rather than delegating broad discretion to subordinates. The motives of a school that fails to do so should be considered suspect.

The use of race in college admissions has been described by UT Austin admissions officers and by the courts at various times as merely a “factor of a factor of a factor,”⁶⁷ a “plus factor,”⁶⁸ and as a factor that can “tip the balance.”⁶⁹ Fortunately there exist data that can test the accuracy of these characterizations. “[I]n data from 40 law schools gathered from their 2005-2007 admissions cycles, the median odds ratio on black compared to white admissions was 150—meaning black applicants had far greater odds of admission than white applicants.”⁷⁰ When one group has 150 times the likelihood of being admitted compared to another group, we suggest that this is not merely a thumb that can tip the balance scale but a bowling ball on the

⁶⁴ *Fisher v. Univ. of Texas*, 758 F.3d at 647: “Given the test score gaps between minority and non-minority applicants, if holistic review was not designed to evaluate each individual’s contributions to UT Austin’s diversity, including those that stem from race, holistic admissions would approach an all-white enterprise.”

⁶⁵ *Students for Fair Admissions, Inc. v. President and Fellows of Harvard College (Harvard Corporation)*, Civil Action No. 1:14-cv-14176-ADB.

⁶⁶ *See Grutter v. Bollinger*, 539 U.S. 306, 328 (2003) (“Today, we hold that the Law School has a compelling interest in attaining a diverse student body.”).

⁶⁷ *Fisher II*, 136 S.Ct. at 2207.

⁶⁸ *Regents of the University of California v. Bakke*, 438 U.S. 265, 312-21 (1978).

⁶⁹ *Id.* at 316.

⁷⁰ Robert Steinbuch, *A Different Take On Why Law Schools Are Not Admitting More Black Students*, NAT’L JURIST, at www.national-jurist.com/national-jurist-magazine/different-take-why-law-schools-are-not-admitting-more-black-students, downloaded 3/2/19.

balance scale.⁷¹ Race is not being used just as a “tie-breaker.” Holistic ratings effectively disguise the true policy. When the Supreme Court opined that “it is not a failure of narrow tailoring for the impact of racial consideration to be minor,”⁷² it was not apprised of the actual quantitative magnitude of the racial consideration. UT failed to do that analysis.

Of course, holistic ratings completely hide the measurement of each of the criteria upon which the holistic rating is based. In his dissent Justice Alito repeatedly criticized UT Austin for not keeping data and also for failing to analyze data they already have in their possession.⁷³ Such data and their analysis would have enabled UT to answer critical questions about whether various diversity factors within the PAS had their intended effect and whether “holistic admittees” had even enrolled in classes lacking in diversity. UT Austin admitted that it does not keep any statistics on the impact of their holistic racial admission system!⁷⁴

B. Proving the Educational Benefits of Diversity

1. Measurement of the purported benefits.

Many of the purported educational benefits of diversity would have been extremely easy to assess. For example, the Supreme Court opined that “student body diversity promotes learning outcomes.”⁷⁵ UT Austin could have given standardized tests to students in multiple instances of the same course but which varied in the diversity of the enrolled students. Both UT and outsiders (like courts) could then compare the knowledge gained in the various instances. Does diversity result in greater gains in learning? UT produced absolutely no evidence whatsoever that its diversity admissions program had fostered the educational benefits that UT had listed as its goals. Perhaps the program had fostered these goals; perhaps it had not. This is a question that could have been easily answered empirically.

In defending its admission policy, UT Austin explained that another benefit of diversity is to foster the “robust exchange of ideas.”⁷⁶ One way to test whether diversity fosters the achievement of this goal is to tape record class sessions which differ in the diversity of the enrolled students. Ask neutral listeners to rate the liveliness of the discussions or the robustness of the exchange of ideas. There already exists research evidence that might bear upon this issue. Many companies are very highly interested in the productivity and creativity of their work teams. The typical research finding on this topic is that diversity of knowledge among team members is

⁷¹ See Peter Arciadiacono & Michael Lovenheim, *Affirmative Action and the Quality-Fit Trade-Off*, 54 J. ECON. LIT. 3, 12-13 (2016) [hereinafter “Arciadiacono & Lovenheim, *Trade-Off*”]; Peter Arciadiacono & Michael Lovenheim, *Affirmative Action In Undergraduate Education*, 7 ANN. REV. ECONS. 487, 495-96 (2015).

⁷² *Fisher II* at 2212.

⁷³ *Id.* at 2216, 2227, 2238 (Alito, J., dissenting).

⁷⁴ *Id.* at 2220.

⁷⁵ *Fisher II*, 136 S.Ct. at 2210, quoting *Grutter*, 539 U.S. at 330.

⁷⁶ *Fisher II* 136 S.Ct. at 2211.

an important facilitative factor but that diversity of ethnicity is not.⁷⁷ Although the same phenomenon would not necessarily be found in college classroom discussions, these research findings do highlight the need of a college to document the benefits it claims in order to prove that its program is “narrowly tailored” and therefore satisfies strict scrutiny.

2. Assessing outcomes vs. self-reports of improvement.

Because self-reports of one’s own improvement are highly suspect,⁷⁸ there have been attempts to overcome this problem by using alternative measurement techniques. The first is the indirect method. In their study Rothman and colleagues asked students, faculty, and administrators from 140 campuses no questions about diversity.⁷⁹ Instead they asked the participants about their own campus’ environment and their own educational experience. These variables were then examined in relation to the enrollment diversity of their campus. The hypothesis to be tested, of course, was that more diversity on a campus would be correlated with positive educational and environmental outcomes. However the results “. . . showed an inverse relationship between enrollment diversity and evaluations of educational quality by students, faculty, and administrators.”⁸⁰ In other words, more diversity was related to *less* educational quality in the view of the students, faculty, and administrators. However “[a]mong faculty and administrators, enrollment diversity was significantly associated with reports of more positive treatment of minority students and less racial discrimination. But these findings were offset by the absence of similar results among students, who also reported more personal victimization as diversity increased.”⁸¹

The most obvious way to avoid the self-report problem is to assess outcomes, not self-reports. This is not a novel suggestion.⁸² Being licensed to practice is a critical outcome measure for aspiring physicians. So Kulatunga-Moruzi and Norman⁵¹ examined the validity of various cognitive and non-cognitive criteria in predicting eventual performance of medical school applicants on the licensing examination of the Medical Council of Canada.⁸³ Both the medical school application and the licensing examination have two components—cognitive and non-

⁷⁷ See Greg L. Stewart, *A Meta-Analytic Review of Relationships Between Team Design Features and Team Performance*, 32 J MGMT. 29 (2006); Sujin K. Horwitz & Irwin B. Horwitz, *The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography*, 33 J MGMT. 987 (2007).

⁷⁸ See *supra* notes 50 and accompanying text.

⁷⁹ Stanley Rothman et al., *Does Enrollment Diversity Improve University Education?*, 15 INT’L J. PUB. OPIN. RESEARCH 8 (2003).

⁸⁰ *Id.* at 16.

⁸¹ *Id.*

⁸² See Justin Pidot, *Intuition or Proof: The Social Science Justification For The Diversity Rationale in Grutter v. Bollinger and Gratz v. Bollinger*, 59 STAN. L. REV. 761, 793 (2006).

⁸³ Chan Kulatunga-Moruzi & Geoffrey R. Norman, *Validity of Admissions Measures in Predicting Performance Outcomes: The Contribution of Cognitive and Non-Cognitive Factors*, 14 TEACH AND LEARN MED. 34 (2002).

cognitive, the latter being sub-divided into measures of communication and problem-solving skills. Having identical components implies that the medical school application and licensing examination are very well aligned. The results of the analyses were that the non-cognitive factors in the application were nearly worthless in predicting anything—cognitive or non-cognitive—on the licensing examination. What saved the non-cognitive factors from a verdict of complete worthlessness was that the personal interview during the application process did predict the performance on the communication portion of the licensing examination. However, the standardized MCAT (Medical College Admissions Test) verbal reasoning score correlated with the communication portion of the licensing examination far better than did the interview score. Also of interest was that the autobiography portion of the admissions process predicted absolutely nothing on the licensing examination. This finding raises questions about the use of the personal essay in race-conscious admissions.

UT Austin and other colleges and universities could use an important outcome measure such as graduation. Once an outcome measure is chosen, officials can then determine how to narrowly tailor the admissions procedure. For example, how does each component of the PAS separately influence the graduation rate? Do those whose PAS is boosted by having spoken a foreign language at home graduate at a higher rate than those whose PAS is not boosted by the presence of this factor? Such an analysis would allow the UT admissions officials to “tailor” the admissions procedure quite well, thus satisfying the Supreme Court’s requirement. UT’s or any other university’s admission committee could then include the factors that predict the outcome measure and discard from consideration the factors that don’t. Simple empirical analysis can answer such important questions, but UT officials seem unmotivated to perform such investigations.⁸⁴ Without empirical analysis, unverified intuition is all that admissions officials can use. That does not satisfy strict scrutiny.

As an example, UT Austin asserted that African-American students from privileged backgrounds “have great potential for serving as a ‘bridge’ in promoting cross-racial understanding, as well as in breaking down racial stereotypes.”⁸⁵ The Texas Top Ten Percent Law grants admission to persons who finished in the top ten percent of their high school graduating class. UT asserted that the minority students admitted in this manner were drawn from high schools the majority of whose students were minority students. UT Austin further asserted that these students often come from poor, disadvantaged families, and would be less able to break down racial stereotypes than would more affluent minorities who could be admitted using the “special circumstances” component of the PAS score.⁸⁶ Is this intuition correct? There are plenty of validated tests of the propensity to stereotype.⁸⁷ Do White students in classes with Top Ten Percent minority students show more stereotyping than do White students in classes

⁸⁴ *Fisher II*, 136 S.Ct. at 2220, 2226 (Alito, J., dissenting).

⁸⁵ *Id.* at 2231 (Alito, J., dissenting).

⁸⁶ *Id.*

⁸⁷ See Joshua Correll et al., *Measuring Prejudice, Stereotypes and Discrimination*, in THE SAGE HANDBOOK OF PREJUDICE, STEREOTYPING AND DISCRIMINATION 45 (John F. Dovidio et al. eds.2010).

with more privileged minority students? These are all empirical questions whose answers should help craft the admissions process.⁸⁸ There is no need to rely on intuition, which may be incorrect.

Not only can the components of the PAS be examined empirically to tailor the admission process, the academic credentials can also be examined in this manner. In his classic text, *The Nature of Prejudice*, Gordon Allport extolled the benefits of inter-racial contact in diminishing stereotypes.⁸⁹ This would seem to be an enormous benefit of diversity in admissions. However Allport cautioned that if the groups were of different status, this benefit would not be achieved and might even reverse.⁹⁰ Consider the fact that the disclosure of the law school data at the University of Michigan revealed that the median African-American admitted student had an academic index at the second percentile of the White distribution.⁹¹ At the University of Wisconsin and the University of Virginia, the median African-American admitted student had an academic index *below* the first percentile of the White distribution.⁹² Thus at these three universities the two racial groups were unequal in the status of their academic credentials, clearly an important factor in their ensuing law school performance.⁹³ Would the presence of such students reduce or exacerbate racial stereotyping?

Some existing evidence suggests that the current practice of affirmative action actually worsens race relations, at least at highly selective schools. A study of 173 colleges and universities by the National Association of Scholars (“NAS”) has found that African-American students at many schools have demanded and received racially separate facilities, organizations, and programs.⁹⁴ This development is the polar opposite of the traditional civil rights movement’s goal of racial integration. The NAS report speculates that this neo-segregation is a response to the mismatch of academic credentials between White and African-American students created by current affirmative action programs.⁹⁵

⁸⁸ See Pidot, *supra* note 82, at 806.

⁸⁹ GORDON W. ALLPORT, *THE NATURE OF PREJUDICE* (1954).

⁹⁰ *Id.* at 274.

⁹¹ See Arcidiacono & Lovenheim, *Trade-Off*, *supra* note 71, at 12.

⁹² *See id.* at 13.

⁹³ See Alex M. Johnson, Jr., *The Destruction of the Holistic Approach to Admissions: The Pernicious Effects of Rankings*, 81 IND. L.J. 313, 313 (2006) (stating that the LSAT correlates .4 with first-year law school grades).

⁹⁴ National Association of Scholars, Report: Neo-Segregation at Yale, at https://www.nas.org/articles/separate_but_equal_again_neo_segregation_at_yale.

⁹⁵ *Id.* See also Peter Arcidiacono et al., *Racial Segregation Patterns in Selective Universities*, 56 J.L. & ECON. 1039, 1039 (2013) (study finding “significant friendship segregation, particularly for blacks” among students at selective universities, “in part because affirmative action results in large differences in the academic backgrounds of students of different races”).

This development also raises questions about UT's claims, accepted by the Court majority in *Fisher II*, that its program was necessary to overcome "feelings of loneliness and isolation" experienced by minority students and to achieve a "critical mass" of minority students.⁹⁶ Although affirmative action increases the number of minority students at a school, the increase does not seem to facilitate the increased integration of courses that UT said the "critical mass" was intended to achieve nor to overcome minority students' "feelings of isolation and loneliness." Rather, schools seem to have resorted to racial separation to address those feelings.

The intuition of the admissions officials about the effects of the affirmative action programs may or may not be supported by more precise empirical data. Those data could provide guidance to such officials concerning the tailoring of the academic index threshold for admission in order to achieve the intended benefits of diversity as well as to satisfy judicial strict scrutiny.

Conclusion

The Court in *Fisher II* said that a university has an "ongoing obligation to engage in constant deliberation and continued reflection regarding its admissions policies."⁹⁷ Holistic admissions programs like those employed by the University of Texas-Austin in *Fisher* are opaque because they hide the significance given to such constitutionally crucial factors as the race of the applicant. In *Fisher* the Supreme Court did not consider this defect because plaintiffs did not raise the issue. In a case where the issue is properly raised, no university could reasonably assert that its holistic admission procedure is "narrowly tailored" so as to satisfy the Court's requirement of strict scrutiny. The influence of any component of such a holistic program would be completely opaque and therefore impossible to scrutinize.

In the university admissions process, even the best predictors bear only a probabilistic relation to the to-be-predicted event. Does coming from a home in which English is not the primary language foster or lessen the probability of graduation? As long as holistic ratings are used, admissions officers will never ascertain what specific factors promote academic success, and the courts will never be able to learn exactly what procedures are used by admissions officers. Assuming that (a) admissions officers do want to know what factors are related to academic success, and (b) courts do want to know factors are influential in admissions officers' decisions, it is extremely easy to use disaggregated ratings to accomplish both goals.

The use of race must be "narrowly tailored to achieve the university's permissible goals."⁹⁸ "A university's goals cannot be elusory or amorphous—they must be sufficiently measurable to permit judicial scrutiny of the policies adopted to reach them."⁹⁹ So Step 1 is to enunciate specific measurable goals. Graduation? GPA? These are examples of the to-be-predicted event or predicted index.

⁹⁶ *Fisher II*, 136 S.Ct. at 2212.

⁹⁷ *Id.* at 2215.

⁹⁸ *Id.* at 2203, citing *Fisher I* at 2419-20.

⁹⁹ *Id.*

Step 2 is *not* to give one overall rating after eyeballing a large number of potential factors. Step 2 should be to rate each and every factor under consideration. The research literature teaches us that humans are extremely poor at the daunting task of combining multiple probabilistic indicators.¹⁰⁰ Employ a multiple regression equation to provide a clear and accurate answer as to what factors should be used and which should be ignored. Predictive accuracy will be enhanced, and the resulting process will be transparent to all. These are worthy goals, and they are easily within our reach. Admissions programs that do not follow such a process are not “narrowly tailored” and therefore should not be held to satisfy strict scrutiny.

¹⁰⁰ Dawes, *supra* note 36, at 1671-72. [8/1/19]