
ARTICLES

SUPPLY SIDE OR DISCRIMINATION? ASSESSING THE ROLE OF UNCONSCIOUS BIAS

*Amy L. Wax**

I. INTRODUCTION

An increasingly prominent theme in legal scholarship is that unconscious bias is an important contributor to social disparities by race and gender. The intensive focus on inadvertent motives is grounded in laboratory observations that purport to demonstrate that people's split-second reactions are influenced by group identity. Such responses, it is argued, routinely fuel discriminatory conduct against women and minorities. It follows that, where disparities exist, they can usually be traced back to the operation of biased psychological processes of which people are largely unaware.

This Article examines these arguments by considering, and critiquing, the claimed link between unconscious bias and social disadvantage. The discussion begins, in Part II, with the recent case of *Wal-Mart Stores, Inc. v. Dukes*,¹ in which the plaintiffs argued that Wal-Mart's treatment of its female employees was uniformly, even if inadvertently, contaminated by gender stereotypes. In examining this contention, Part II elaborates on the distinction between risk and discrimination—between the *potential* for gender or other forbidden factors to influence decisions versus whether those factors are *actually* taken into account. Part III reviews, and finds wanting, the scientific evidence that claims to attribute real-world discriminatory treatment, and the resulting adverse outcomes, to unconsciously biased motives. Part IV.A discusses the general difficulties of demonstrating discrimination, which requires distinguishing between unlawful motives (conscious or unconscious) and other factors as a cause of unfavorable treatment. Part IV.B discusses the pervasive legal assumption that outcome disparities reflect discriminatory treatment and the documented intergroup differences that undermine that approach. Part V looks at measures proposed to address inadvertent bias, and notes the lack of evidence that any of them will have the intended effect of diminishing unlawful discrimination. Parts VI and VII argue that, given the existence of significant "supply side" factors that explain persistent intergroup differences, it is unlikely that unconscious discrimination is an important source of social disadvantage, especially for race. The discussion then concludes with the suggestion that the roots of

* Robert Mundheim Professor of Law, University of Pennsylvania Law School. This Article is based on a presentation given at *Temple Law Review's* November 2011 Symposium: "The Evolution of Civil Rights Litigation: Using Social Science and Statistics to Prove Employment Discrimination and Predatory Lending."

1. 131 S. Ct. 2541 (2011).

inequality are best addressed directly through policies aimed at their alleviation rather than through anti-discrimination law.

II. THE *WAL-MART* CLASS ACTION: POSSIBLE VERSUS ACTUAL STEREOTYPING

In *Dukes v. Wal-Mart Stores, Inc.*,² women employed at Wal-Mart stores nationwide sued the giant retailer, claiming that female workers as a class received less favorable treatment than males.³ Specifically, the plaintiffs alleged that women “are paid less than men in comparable positions,” despite their better performance and greater seniority, and that women “receive fewer—and wait longer for—promotions to . . . management positions than men.”⁴

Central to the plaintiffs’ case was the allegation of stereotyping.⁵ According to the plaintiffs’ theory, Wal-Mart supervisors harbored a set of ideas and expectations about the roles of women (and men).⁶ These ideas and expectations swayed personnel decisions in ways that systematically disadvantaged female workers. A linchpin of this alleged scenario is the exercise of managerial discretion. The plaintiffs contended that managers’ unguided authority to promote and reward employees at the store level allowed stereotypes to influence how workers were treated.⁷ Senior management did nothing to correct these patterns, and indeed sanctioned the discretionary methods that gave stereotypes free rein.⁸

Although the allegations in *Wal-Mart* did not rest solely on claims of unconscious bias,⁹ they were consistent with the possibility that Wal-Mart’s unfavorable treatment of women was influenced by inadvertent motives. Indeed, the notion that unconscious stereotyping was operating at Wal-Mart fits with the type of arguments that have recently been made against large national corporations. Wal-Mart staunchly professed a commitment to equal opportunity in its hiring and promotional practices.¹⁰ The company denied any hint of an official policy to treat male and female employees differently, and no finding in the case contradicted the company’s commitment to evenhanded procedures. The plaintiffs nonetheless maintained that the company regularly made biased choices. Their theory of the case boiled down to the allegation that stereotypes of women—as, for example, less capable, reliable, or committed—

2. 603 F.3d 571 (2010), *rev’d*, 131 S.Ct. 2547 (2011).

3. *Wal-Mart*, 603 F.3d at 577.

4. *Id.*

5. See *Dukes v. Wal-Mart Stores, Inc.*, 222 F.R.D. 137, 145 (N.D. Cal. 2004) (“Plaintiffs submit evidence that Wal-Mart cultivates and maintains a strong corporate culture which includes gender stereotyping.”).

6. *Id.* at 165–66.

7. *Id.* at 148–51.

8. *Id.* at 145; see also Lesley Wexler, *Wal-Mart Matters*, 46 WAKE FOREST L. REV. 95, 105–11 (2011) (providing detailed explanation of plaintiffs’ theory on discrimination).

9. See *Wal-Mart*, 222 F.R.D. at 154–66 (discussing plaintiffs’ statistical evidence of gender disparities among Wal-Mart’s workforce and anecdotal evidence of discriminatory intent by some Wal-Mart managers).

10. See Brief for Petitioner at 3, *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541 (2011) (No. 10-277), 2011 WL 201045 (“Wal-Mart’s company-wide policy . . . expressly bars discrimination based on sex, and Wal-Mart has consistently promulgated and enforced equal opportunity policies . . . prohibit[ing] unlawful discrimination.”).

were allowed to infect employment decisions.¹¹ Although plaintiffs' arguments did not expressly rest on the assertion that stereotypes operated inadvertently, their contentions were consistent with supervisors and managers being unaware that their decisions were distorted by improper stereotypes. In short, according to plaintiffs, the supervisors' good faith did not rule out liability.¹²

Critical to the plaintiffs' case in *Wal-Mart* was the testimony of William Bielby, an expert who frequently participates in bias claims against large corporate employers.¹³ Central to Bielby's testimony was the assertion that "[p]ersonnel policy and practice at Wal-Mart as implemented in the field has features known to be vulnerable to gender bias."¹⁴ In support of this, Bielby recounted the record evidence that individual store managers exercised broad discretion in the promotion and pay of store employees. According to Bielby, "there are no written criteria for selecting hourly associates for promotion into management . . . beyond the minimum requirements."¹⁵ Therefore, managers "have discretion to devise their own criteria, with no monitoring or oversight over how those criteria are devised or applied."¹⁶ In Bielby's view, such unguided discretion posed a significant risk that gender stereotypes could influence personnel decisions and thereby operate to the disadvantage of female employees.¹⁷ In justifying his conclusion that Wal-Mart's employment practices were vulnerable to bias, Bielby cited "[a] large body of research" demonstrating that "the tendency to invoke gender stereotypes in making judgments about people is spontaneous and automatic."¹⁸ He noted that "[a]s a result, people are often unaware of how stereotypes affect their perceptions and behavior, and individuals whose personal beliefs are relatively free of prejudice or bias are susceptible to stereotypes in the same ways as people who hold a personal animosity towards a social group."¹⁹

The Bielby testimony has been the subject of considerable controversy. Some social scientists—including leading proponents of the "social framework" approach that Bielby purported to apply in his Wal-Mart statements—have claimed that Bielby's allegations represent a misuse of the research upon which he relies.²⁰ Although a full analysis of the controversy is not possible here, one key point is worth considering in greater detail.

11. For a review of the arguments in *Wal-Mart*, see Wexler, *supra* note 8, at 105–11.

12. See Amy L. Wax, *The Discriminating Mind: Define It, Prove It*, 40 CONN. L. REV. 979, 987 (2008) (noting allegations of discrimination against employers with professed commitment to equal opportunity).

13. Grace E. Speights & Bernard R. Siskin, *The Impact of Statistical Evidence in Class Action Litigation*, SS032 ALI-ABA 859, 883 (2011) ("Dr. William Bielby has clearly been the most prominent proponent of the stereotyping theory and has testified regarding the theory in many cases.").

14. Declaration of William T. Bielby, Ph.D in Support of Plaintiffs' Motion for Class Certification at 5, *Wal-Mart*, 222 F.R.D. 137 (No. C-01-2252), 2003 WL 24571701.

15. *Id.* at 22.

16. *Id.*

17. *Id.* at 19–20 ("In such settings, stereotypes can bias assessments of a woman's qualifications, contributions, and advancement potential, because perceptions are shaped by stereotypical beliefs about women generally, not by the actual skills and accomplishments of the person as an individual.").

18. *Id.* at 17–18.

19. *Id.* at 18.

20. For a discussion of this controversy, see Adam Liptak, *Supreme Court to Weigh Sociology Issue in Wal-Mart Discrimination Case*, N.Y. TIMES, Mar. 27, 2011, at A17.

The *Wal-Mart* litigation was based on the assertion that the procedures used at Wal-Mart were vulnerable to distortion by gender stereotyping. But what does that mean for individual female employees within the class? A close reading of Bielby's testimony reveals that Wal-Mart's practices create *the potential* for decisions to be tainted by illicit considerations of gender instead of resting solely on meritocratic factors such as "qualifications, contributions, and advancement potential." But that is not the same as proving that stereotyping actually influenced any particular employment decision. That distinction turns out to be critical to the procedural issue of class certification that was recently considered by the Supreme Court,²¹ as well as to the underlying merits of the allegations of gender discrimination. Specifically, it is relevant to whether the plaintiff class was sufficiently similar—that is, whether its members had enough in common—to warrant class certification. And, as noted in the Court's decision, the distinction between potential and actual discrimination was pertinent to which individual class members would ultimately be entitled to monetary relief for the harms they suffered from unlawful discrimination.²²

During the oral argument before the Supreme Court on class certification, the attorney for Wal-Mart noted that Bielby maintained only that Wal-Mart's policies created the *potential* for stereotyping. He stressed that Bielby did not purport to identify which particular decisions were biased, or to demonstrate how many decisions were impermissibly influenced or distorted by gender concerns. Indeed, Bielby admitted under questioning that he could not be sure that *any* personnel decisions at Wal-Mart had been swayed by stereotypes. As Wal-Mart's counsel stated, "the sociologist here [Dr. Bielby], who is the glue that's supposed to hold this class together, said he couldn't tell if stereotyping was occurring one half of 1 percent or 95 percent or at all."²³ The attorney for Wal-Mart went on to observe that "[t]he question here is whether . . . we can assume that every decisionmaker acted in the same manner in a way that had . . . the same injury."²⁴ He argued that the answer to this question was clearly no, adding that "the plaintiffs need[] to come forward with something that show[s] that there was this miraculous recurrence at every decision across every store of stereotyping, and the evidence simply doesn't show that."²⁵

Justice Ginsberg appeared to echo these concerns when she suggested at oral argument that Bielby's concessions would likely have some bearing on one aspect of the case—the relief that individual female employees would receive if the class action went forward and Wal-Mart was ultimately found liable for violating Title VII.²⁶ Along these lines, Ginsberg noted a "very serious problem in this case" of "how [to] work out the back pay" due each member of the class.²⁷ Determining back pay requires identifying individuals who actually suffered a loss from unlawful discrimination. The

21. See generally *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541 (2011).

22. *Id.* at 2557–59.

23. Transcript of Oral Argument at 13, *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541 (2011) (No. 10-277).

24. *Id.* at 13–14.

25. *Id.* at 9.

26. *Id.* at 33–34.

27. *Id.* at 33.

need to determine who was victimized, and to measure how much each victim was harmed, highlights the gap between the risk of discrimination (which is alleged to be created by a discretionary evaluation system), and whether, and to what extent, discrimination actually occurred. Ginsberg notes that, because the plaintiff class is very large, “there’s no way [the judge] could possibly try each of these individual[] [cases].”²⁸ Yet the proper determination of the remedy requires just such an individualized inquiry.

That a hearing on each employee is necessary, at least at the remedial phase, suggests that it is impossible to assume on the facts as stated (at least at the preliminary stage of class certification) that every member of the plaintiff class was the victim of unlawful discrimination. Some women at Wal-Mart may not have been harmed by stereotyping or bias. Perhaps many were not. Indeed, some of the women received promotions and pay raises, and many stores were headed by female managers. Furthermore, it cannot be assumed that all of the women who were denied benefits or promotions were the targets of unlawful treatment. Not everyone who does poorly in the workplace is the victim of discrimination. Some negative outcomes—including denials of promotions or raises, discipline, or termination—might be deserved. Even if undeserved, these actions might have nothing to do with gender stereotyping or sex.

All these concerns are in fact reflected in the Supreme Court’s decision in the *Wal-Mart* case, which was issued at the end of the 2010-2011 term.²⁹ The majority opinion resolved the preliminary procedural question of whether the class action was properly certified in the case.³⁰ The Court’s resolution of that issue in *Wal-Mart*’s favor rested in large part on the diversity of circumstances for the individuals in the class and the difficulty of proving actual discrimination against the group. The majority opinion focused on the distinction between the danger of discriminatory stereotyping allegedly presented by managerial discretion, and whether or not the exercise of that discretion actually resulted in gender-based discrimination against particular female employees. As the Court stated, “the recognition that this type of Title VII claim ‘can’ exist does not lead to the conclusion that every employee in a company using a system of discretion has such a claim in common.”³¹ Specifically, even if the plaintiffs could show a general vulnerability, or “pattern and practice” of decisionmaking that posed a heightened risk of discrimination, only class members who actually suffered discrimination would be entitled to back pay.³² That entitlement—which requires a showing of unlawful discrimination towards a particular person—would have to be determined for each employee and would require “additional proceedings.”³³

The case against *Wal-Mart* was grounded in an allegation of disparate treatment—the claim that employment decisions were influenced by the forbidden factor of the employee’s gender. The assertion that gender somehow “caused” some unknown number of unfavorable outcomes for *Wal-Mart*’s female employees lay at the heart of

28. *Id.* at 34.

29. *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541 (2011).

30. *Id.* at 2547.

31. *Id.* at 2554.

32. *Id.* at 2560.

33. *Id.* at 2561.

the plaintiffs' claims in the case. But the Supreme Court's decision makes clear that the distinction between the *possibility* of discrimination and *actual* discrimination cannot be ignored. Plaintiffs' theory of the case turned on the identification of a risk—specifically, the risk that stereotypical assumptions might influence how managers treat women. That risk might be great or small in magnitude; it might apply to a few women, or most, or all. But creating a potential risk is not the same as engaging in unlawful discrimination towards one or more individuals. A risk is only a possibility, and possibilities might not be realized in every case.

Critics will contend that this analysis is churlish and overly simple, because it slights the centrality of the risk of biased employment decisions and the vital importance of reducing or eliminating that risk. The *Wal-Mart* allegation is that certain practices and workplace protocols enhance the possibility that female workers will be stereotyped and that their treatment will be tainted by unlawful concerns. As an institutional matter, this is not a trivial accusation, and such an effect might actually exist. In fact, there are special types of anti-discrimination lawsuits that seek to address such possibilities. So-called “pattern or practice” litigation, although not currently commonplace, seems designed to do this.³⁴ Although the doctrine is complicated, pattern or practice litigation starts with the fact of group disparities, identifies a practice that might be causing them, and then shifts the burden to employers to offer a non-discriminatory explanation for the disparate pattern of outcomes observed.³⁵ If no plausible explanation is forthcoming, courts can “reasonably conclude that the statistically significant disparity is the product of intentional discrimination.”³⁶

Pattern or practice lawsuits infrequently succeed, in part because “plain[] and open[]” discrimination has faded from the landscape,³⁷ and also because defendants have become more sophisticated about asserting persuasive non-discriminatory reasons for unfavorable treatment and group disparities. Likewise, courts have become more skeptical of discrimination as the explanation for observed group outcome differences, and are more aware of plausible alternatives.³⁸ The situation poses even more challenges for plaintiffs when the ultimate goal of litigation is not just to identify and eliminate a tainted or risky practice, but to vindicate individual claims of discrimination. As already discussed, unless a victim has been personally harmed by a forbidden practice, he or she is not entitled to relief. Thus, when employees seek individual redress and compensation, rather than just prospective company-wide reform, the courts cannot avoid the question of individual causation. In awarding a monetary remedy—the type that, understandably, many plaintiffs fervently desire—this problem looms large.

34. For a recent discussion of the Supreme Court's pattern or practice jurisprudence, see Michael Selmi, *Theorizing Systemic Disparate Treatment Law*, 32 BERKELEY J. EMP. & LAB. L. (forthcoming 2011), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1809944.

35. See *id.* (manuscript at 3–6) (describing characteristics of pattern or practice claims as distinguished from disparate impact claims).

36. *Id.* (manuscript at 6).

37. *Id.* (manuscript at 7).

38. *Id.* (noting that it has become “less clear, at least to the courts, that discrimination always provided the best explanation for the observed disparities”). For more on alternative explanations, see *infra* Part III.

III. EXPLANATIONS FOR DISPARATE OUTCOMES: DOES IMPLICIT BIAS PREDICT DISCRIMINATION?

The distinction between risk and discrimination—between the potential for forbidden factors to contribute to an unfavorable decision, as opposed to whether those factors actually do so—is particularly critical to the debate surrounding the role of unconscious bias, and what (if anything) should be done about it. Specifically, the issue is relevant to whether unconsciously biased thinking actually produces discriminatory harms. Resolving this question highlights the importance of considering and ruling out alternative explanations for individual and group-level disparities in outcomes on the job and in other areas of social life. It also requires a critical look at the research that purports to detect unconscious bias, and at the studies that claim to link this bias to discriminatory treatment in the real world.

In the few years since I last addressed these topics,³⁹ many articles have appeared touting the significance of unconscious biases and inadvertent stereotyping for decisionmaking in a range of contexts. While some researchers have cast a rigorous and critical eye on the data and have expressed grave doubt about the value of correlations and predictions,⁴⁰ others have made sweeping claims about the science's far-reaching significance.⁴¹ Unfortunately, too much of this work still fails adequately to grapple with the issues of proof and causation that are central to any proper analysis of the role of unconscious processes in the law of discrimination.

An important research review by Hart Blanton and colleagues, which summarizes and analyzes the social science evidence on unconscious bias and its predictive value, highlights the shortcomings (and misuse) of much of the data in this area.⁴² The focus of this review is on the so-called implicit association test (IAT), a much-studied technique that purports to gauge prejudice towards social groups based on split-second reaction times measured in the laboratory. As developed by Mahzarin Banaji and her colleagues at Harvard, the IAT measures the speed with which subjects associate negative versus positive words or concepts with images of people from different backgrounds.⁴³ The association is thought to reveal unconscious processes similar to stereotyping. In the context of race, for instance, the contention is that people who more quickly link negative ideas to black faces and positive ideas to white faces show that they harbor unconscious negative stereotypes or biases against blacks. Researchers have maintained that people with such negative associations will be more likely to engage in racial discrimination.⁴⁴

39. See generally Wax, *supra* note 12; Amy L. Wax, *Discrimination as Accident*, 74 IND. L.J. 1129 (1999).

40. E.g., Hart Blanton et al., *Strong Claims and Weak Evidence: Reassessing the Predictive Validity of the IAT*, 94 J. APPLIED PSYCHOL. 567 (2009).

41. E.g., John T. Jost et al., *The Existence of Implicit Bias Is Beyond a Reasonable Doubt: A Refutation of Ideological and Methodological Objections and Executive Summary of Ten Studies That No Manager Should Ignore*, 29 RES. ORGANIZATIONAL BEHAV. 39 (2009).

42. See generally Blanton et al., *supra* note 40.

43. See generally Brian A. Nosek et al., *The Implicit Association Test at Age 7: A Methodological and Conceptual Review*, in SOCIAL PSYCHOLOGY AND THE UNCONSCIOUS: THE AUTOMATICITY OF HIGHER MENTAL PROCESSES 265 (John A. Bargh ed. 2007).

44. For a more in depth discussion, see Wax, *supra* note 12, at 1018–21.

The Blanton review concludes that the research permits neither the prediction nor the identification of discriminatory conduct at the individual level.⁴⁵ The lack of credible evidence on the predictive value of measures of unconscious bias is especially acute in the context of race. Because race is an important focus of discrimination law and scholarship, claims regarding the operational role and real-world significance of unconscious racial bias loom large. Employment discrimination is another area of frequent litigation and obsessive concern by legal scholars, with much of the discussion of unconscious bias centering on allegations of unfair treatment in the workplace. Based on the contention that unconscious stereotyping is pervasive in the employment arena, far-reaching proposals for legal and structural reform are now a staple of legal scholarship.⁴⁶ These recommendations, and the evidence claimed to support them, do not withstand scrutiny. Put simply, existing research on unconscious bias does not permit an inference of unlawful racial discrimination in any particular case. In particular, the IAT is of no help, as of yet, in determining which, if any, adverse outcomes for minorities, either in the workplace or elsewhere, are due to racial discrimination as opposed to other forces or factors.⁴⁷ Finally, the paucity of credible evidence extends to measures that have been proposed to alleviate the effects of unconscious bias. As the discussion below reveals, there is no reason to believe that the sweeping reforms recommended by scholars will reduce inadvertent discrimination in the workplace.

The failure to establish a credible connection between implicit-bias evidence and legally actionable discrimination remains a feature of both the social science literature and legal scholarship. Despite extravagant claims to the contrary, recent articles on this topic reveal the paucity of reliable data linking laboratory measures of unconscious bias to real-world instances of discriminatory treatment. Some of the key articles that exemplify this failure are worth examining in more detail.

A recent lengthy review by psychologist John T. Jost, of New York University, and his colleagues is typical in raising expectations that are ultimately unfulfilled.⁴⁸ The Jost article, which appeared in a specialty journal of organizational psychology, reviews reports of IAT results for a range of subjects under a variety of conditions. According to the authors, the research in this area reveals people's pervasive, albeit variable, tendency to associate negative words and characteristics with racial minorities.⁴⁹ That is, there is ample evidence of "positive" IAT results. The authors consider the question of whether these IAT findings "reliably predict[] class-wide discrimination"⁵⁰—that is, whether these measures predict trait-based decisions in the real world.

The decisions of greatest interest—at least to lawyers—are those that potentially give rise to legal action. Consider once again the subject of racial discrimination in

45. Blanton et al., *supra* note 40, at 578.

46. See *infra* Part VI for a discussion of reforms that have been proposed.

47. See *infra* notes 48–85 and accompanying text for a critique of research purportedly demonstrating the predictive value of the IAT.

48. Jost et al., *supra* note 41.

49. *Id.* at 42–46.

50. *Id.* at 46.

employment. The relevant evidence in Jost's review includes only two studies that purport to show a link between implicit bias as measured by the IAT and workplace discrimination.⁵¹ It is striking that, in a twenty-two page, single-spaced review, these two studies represent the sum total of research directly pertinent to the sweeping claim that the IAT "predicts socially and organizationally significant behavior, including [in] employment."⁵² One of the cited studies, describing research conducted in Sweden, claims to establish a correlation between managers' IAT scores and their willingness to interview Swedish or Arab job candidates with otherwise similar job resumes.⁵³ The other cited study is a brief paper in the *American Economic Review* (*AER*) that, according to Jost and colleagues, purports to find that "scores on an implicit stereotyping task involving race and intelligence were correlated with students' likelihood of selecting resumes with African American names, especially among participants who felt rushed while completing a resume-selection task."⁵⁴ However, the referenced paper falls far short of delivering on the promised description.

The purpose of the cited paper, which was designed as a preliminary study and brief introduction to the field, is to familiarize economists with the idea that discrimination "may be *unintentional* and outside the discriminator's awareness."⁵⁵ The *AER* paper includes a cursory, conclusory review of the basics of IAT research, which briefly touches on the question of whether and how strongly the data establish a link between tests of unconscious bias (specifically, the IAT) and real-world behavior. Along these lines, the *AER* paper contains one short (i.e., two paragraph) description of a small pilot project, which was designed to follow up on a previous experimental observation that hypothetical job candidates with black-sounding names were hired less often than candidates with similar resumes and white-sounding names.⁵⁶ Based on a pilot screening of 50 resumes by 115 participating subjects, the *AER* report states that participants who implicitly associated blacks with lower intelligence (based on a rapid link between the words "black" and "dumb") picked relatively fewer black-sounding candidates, especially under time-pressured conditions.⁵⁷ However, no actual numbers were presented and no data analysis was included. The *AER* authors stressed the limited nature and shortcomings of their findings.⁵⁸ They stated, without further explanation, that "we did not find discrimination, on average, in the lab"; that the lab exercise lacked "external validity"; and that the tests faced "implementation problems."⁵⁹

These findings amount to very thin gruel on which to base claims of pervasive unconscious bias in the employment context. The only other studies cited in the Jost review focus on behaviors that are far removed from the actual employment context,

51. *Id.* at 49 (citing Marianne Bertrand et al., *Implicit Discrimination*, 95 AM. ECON. REV. 94 (2005); Dan-Olof Rooth *Implicit Discrimination in Hiring: Real World Evidence* (Inst. for the Study of Labor (IZA), Discussion Paper No. 2764, Apr. 2007)).

52. *Id.* at 39 (abstract).

53. *Id.* at 49 (citing Rooth, *supra* note 51).

54. *Id.* (citing Bertrand et al., *supra* note 51).

55. Bertrand et al., *supra* note 51, at 94.

56. *Id.* at 96-97.

57. *Id.*

58. *Id.* at 97.

59. *Id.*

and most do not even deal with race. For example, the reports connect measured implicit reaction times to the tendency to engage in binge drinking,⁶⁰ to choose one of two candidates for the President of Italy,⁶¹ to express sexual attraction towards children,⁶² and to suffer from suicidal ideation.⁶³ The point of these studies is to contrast express attitudes with implicit associations as predictors of actual behavior. The goal is to try to demonstrate that implicit attitudes are a better guide to future behavior than what subjects actually say about their preferences or what they themselves predict they will do. But that is a far cry from demonstrating that the IAT can predict who will discriminate against workers based on race.

The only other research paper Jost and colleagues rely on to demonstrate the role of implicit *racial* bias in predicting real-world decisions is one claiming to show that medical trainees with higher (more “anti-black”) IAT scores are more likely to recommend thrombolytic-type therapy for white rather than black patients described in hypothetical profiles that suggest a diagnosis of coronary-artery disease.⁶⁴ Jost and colleagues describe this study as demonstrating “that implicit racial bias predicts the withholding of valuable medical treatment from some patients.”⁶⁵ This allegedly suggests “that one’s degree of implicit bias can have life-or-death consequences for others.”⁶⁶ However, the dire assertion that implicit racial bias leads to life-threatening deprivation of treatment represents a significant overreading of the study. The reported study suffers from a critical flaw: there is no evidence presented that the particular therapies prescribed on the basis of the hypothetical profiles would lead to worse outcomes for the black patients.⁶⁷ Without that direct evidence of harm, there is no reason to believe that “more is better.” Accordingly, there is no basis for inferring the general proposition that blacks will suffer from the physician behaviors evinced by the study or that doctors who prescribe thrombolytic therapy less frequently to blacks are biased “against” that group. In addition, there is no *a priori* reason to assume that, in the actual world of medical practice, blacks and whites with similar symptoms should necessarily receive the same treatment. The success and appropriateness of a particular therapy may not just be a function of the patient’s specific symptoms and medical diagnosis, but could also be influenced by genetic factors, general health status, and surrounding social circumstances that might bear on the effectiveness and response to

60. Jost et al., *supra* note 41, at 52.

61. *Id.*

62. *Id.* at 53.

63. *Id.*

64. See Alexander R. Green et al., *Implicit Bias Among Physicians and Its Prediction of Thrombolysis Decisions for Black and White Patients*, 22 J. GEN. INTERNAL MED. 1231 (2007), discussed in Jost et al., *supra* note 41, at 51–52.

65. Jost et al., *supra* note 41, at 52.

66. *Id.*

67. *Id.* at 51–52. The Jost review also present studies that purport to show the existence of implicit biases directly (without any correlation with IAT scores). For example, some research claims that police recruits, in simulated confrontations, are more likely to shoot black than white subjects. *Id.* at 50–51. However, the data also indicate that these effects are small and short-lived and wash out with training. *Id.* at 51.

different therapies.⁶⁸ To the extent that these might systematically differ by race, it is at least arguable that doctors might end up making different decisions, at least on average, for patients from different groups, and that those patterns might benefit patients overall and thereby be medically justified. In sum, although doctors might sometimes end up treating races differently, or even taking race into account in some medical contexts, it is far from clear that this will prove detrimental. To be sure, the role of race in medical treatment is controversial.⁶⁹ But that controversy does not establish proof of harm.

In sum, despite the virtually non-existent evidence to support their sweeping conclusions, Jost and co-authors assert with vigor and confidence that implicit bias pervades the business world, and that minorities are routinely harmed by its operation.⁷⁰ As a measure of their desperation and the paucity of hard facts, the authors ultimately abandon any attempt to demonstrate the predictive power of the IAT with actual data, and simply shift the burden of proof. Citing a number of audit studies purporting to show that “racial discrimination continues to exist in the marketplace,” but that have nothing to do with measures of implicit bias, the authors simply state that “no research exists to sustain [the] buoyant assumption that managers would be immune to the effects of implicit bias.”⁷¹ This bald assertion, which runs contrary to the entire structure of anti-discrimination law and the burden of proof for civil liability, suggests that employers are guilty until proven innocent and that plaintiffs in discrimination lawsuits are always right. In effect, the statement represents an attempt to short-circuit the problem of prediction entirely. In inviting legal plaintiffs to dispense with the arduous work of establishing the relationship between laboratory research on implicit bias and actual unlawful discrimination, the statement amounts to nothing more than special pleading, wishful thinking, and hand waving. Yet this move is not exceptional. Unsupported assumptions are all too common, not just in the research literature but also in the legal scholarship that draws upon it.

A lengthy law review article,⁷² published in 2006 and written by a leading IAT researcher, Mahzarin Banaji, and a UCLA law professor, Jerry Kang, is typical of the genre. This piece boldly asserts in its introduction that “the magnitude of implicit bias towards members of outgroups or disadvantaged groups is large.”⁷³ It also states that “implicit bias influences evaluations of and behavior towards those who are the subject of the bias.”⁷⁴ Based on these statements, the authors then proceed to justify affirmative

68. See, e.g., Esteban González Burchard et al., *The Importance of Race and Ethnic Background in Biomedical Research and Clinical Practice*, 348 NEW ENG. J. MED. 1170, 1174–75 (2003) (discussing need to further understand “variations among racial and ethnic groups in the prevalence and severity of diseases and in responses to treatment” (emphasis added)).

69. See, e.g., *id.* at 1171 (suggesting that “[e]xcessive focus on racial or ethnic differences” may create “risk of undervaluing the great diversity that exists among persons within groups”).

70. Jost et al., *supra* note 41, at 46–51 (describing “ten studies no manager should ignore”).

71. *Id.* at 49.

72. Jerry Kang & Mahzarin R. Banaji, *Fair Measures: A Behavioral Realist Revision of “Affirmative Action,”* 94 CAL. L. REV. 1063, 1064 (2006).

73. *Id.* at 1064.

74. *Id.* For a more recent article, also co-authored by Jerry Kang, that makes similar points, see Jerry Kang & Kristin Lane, *Seeing Through Colorblindness: Implicit Bias and the Law*, 58 UCLA L. REV. 465 (2010).

action in a host of domains not as a historical corrective but as a “response[] to discrimination in the here and now.”⁷⁵ Because unconscious bias is pervasive and significant, say Kang and Banaji, racial preferences should be adopted until such time as “measures of implicit bias for a region or nation are at zero or some rough behavioral equivalent.”⁷⁶ In other words, the results of IAT research justify the indefinite continuation of racial affirmative action across the board.

On the critical question of whether IAT scores actually correlate with real-world behaviors, Kang and Banaji rely primarily on the results of an unpublished meta-review (which was subsequently published in 2009), that analyzed 224 claimed IAT-behavioral correlations reported in 21 published and 31 unpublished papers.⁷⁷ Kang and Banaji describe this meta-analysis as showing that “implicit biases correlate[] with real-world behaviors like being friendly toward a target, allocating resources to minority organizations, and evaluating job candidates.”⁷⁸ However, the authors themselves concede that most of the behaviors that were actually measured—such as eye contact and “stiff body language”—are only “intermediary steps to some final decision” and not the final decision itself.⁷⁹ Indeed, the only cited study that purported to establish a direct connection between measures of implicit racial bias and a bottom-line choice regarding job candidates of different races showed unimpressive results: the study was small, and a significant IAT-outcome correlation was measured only when “the president of the firm signaled his preference [to the study subjects] for a White hire.”⁸⁰ In other words, the study subjects could only be observed to engage in measurable discrimination when explicitly directed to do so by their boss! It is worth noting that this situation differs markedly from the typical scenario, as exemplified in *Wal-Mart*, where a business vigorously embraces equal opportunity and non-discrimination, but workers contend that subconscious bias nonetheless undermines the operation of those principles.⁸¹

Finally, Kang and Banaji make much of the paper by Green et al.—cited by Jost and colleagues and described above—that purports to link medical interns’ IAT scores to their treatment decisions for black and white patients with possible coronary artery disease.⁸² In describing the study, Kang and Banaji report that the physicians prescribed appropriate state-of-the-art thrombolytic therapy with similar frequency to

75. Kang & Banaji, *supra* note 72, at 1065.

76. *Id.* at 1066.

77. See Anthony G. Greenwald et al., *Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity*, 97 J. PERSONALITY & SOC. PSYCHOL. 17 (2009), cited in Kang & Banaji, *supra* note 72, at 1073.

78. Kang & Banaji, *supra* note 72, at 1073.

79. *Id.*; see also Kang & Lane, *supra* note 74, at 488 (noting that research attempting to correlate implicit biases with behavior is focused on a range of responses including intermediate variables such as “non-verbal behavior, social judgments, physiological responses”).

80. Kang & Banaji, *supra* note 72, at 1073–74.

81. *Id.* at 1074 (citing Jonathan C. Ziegert & Paul J. Hanges, *Employment Discrimination: The Role of Implicit Attitudes, Motivation, and a Climate for Racial Bias*, 90 J. APPLIED PSYCHOL. 553, 556 (2005)). The Zeigert study has been vigorously criticized as unsound by Blanton and colleagues based on their review of the data underlying the report. Blanton et al., *supra* note 40, at 578–80.

82. See Kang & Banaji, *supra* note 72, at 1074–75 (citing Green et al., *supra* note 64).

both black and white hypothetical patients, but that the black patients were more likely to receive a diagnosis of CAD.⁸³ In their view, this “creat[es] a greater discrepancy between diagnosis and treatment for black than white patients.”⁸⁴ The authors construe the study as revealing the doctors’ adoption of a “stereotype that blacks were stubborn and noncompliant and therefore likely to refuse treatment,” but they cite no direct evidence for this interpretation.⁸⁵ Once again, the authors say nothing about the health outcomes that might result from the treatment decisions at issue, which were based on short hypothetical vignettes rather than flesh and blood patients. The authors do not discuss whether blacks in the real world would suffer adverse effects relative to whites from somewhat lower rates of treatment with a particular type of therapy, or whether that question might be dependent on broader background or contextual details that were not revealed in the study. In any event, whether “withholding” certain treatments, or administering others, is helpful or hurtful to individuals or groups is not amenable to theoretical speculation from a small number of short hypothetical descriptions. Rather, the issue is an empirical one, yielding only to extensive, painstaking, and disinterested data collection and analysis.

IV. DISPARATE OUTCOMES: DISCRIMINATION, OR SOME OTHER EXPLANATION?

A. *Leaping to Conclusions on Discrimination*

The articles discussed above share a crucial defect: they assume background facts that need to be demonstrated and thereby leap to a host of unsupported conclusions. The authors take for granted that real-world decisions are regularly and predictably distorted by forbidden and unconscious biases and that outcome differences are the product of unlawful discrimination. This obviates the need to demonstrate actual discrimination in particular cases. Causation is assumed, and alternative explanations disregarded or dismissed.⁸⁶

Paradoxically, one claim that appears in the literature is that the degree of bias revealed by the IAT is *predictive* of discrimination.⁸⁷ That is, there is a quantitative correlation between implicit reactions and real-world behavior. This assertion is at odds with another oft-repeated refrain, which is that both bias and discrimination are *constant* and *pervasive*—a background feature of the social landscape. The problem is that IAT test scores show considerable variation. If the IAT is a predictor, this would imply that the degree of implicit bias—and the resulting discrimination—are highly variable as well, with some people discriminating very little or not at all. But that negates the vision of bias as infecting all social situations and contaminating all decisionmaking. At the very least, discrimination is a matter of degree. It follows that,

83. *Id.* at 1074.

84. *Id.*

85. *Id.*

86. *See generally* Wax, *supra* note 12 (discussing importance of causation in discrimination claims).

87. *See, e.g.,* Green et al., *supra* note 64, 1237–38 (claiming that IAT scores predict the tendency to underprescribe thrombolytic therapies to black patients); Jost et al., *supra* note 41, at 47–48 (listing studies that claim IAT scores correlate with probability of displaying certain biased behaviors).

in some circumstances, discrimination will be minimal or non-existent. The challenge, therefore, is to figure out when discrimination operates and when it does not. It cannot be assumed to occur at all times, or even most of the time, when disparities by race and gender exist. But figuring out when unlawful discrimination is the cause of disparate outcomes has been the task of anti-discrimination law from the beginning.

Lurking behind the talk of unconscious bias is the need to explain adverse treatment. With respect to employment, which is the focus of the discussion here, allegations of unlawful discrimination arise when a person is the target of an unfavorable employment decision, such as the failure to be hired or promoted into a desired position, or the denial of sought-after compensation or benefits. Is the outcome due to unlawful discrimination, or does it have some other explanation? This is the critical question in the down-to-earth, real-world context in which allegations of discrimination arise. Some employees are feckless, lazy, or incompetent. Others are simply less proficient, capable, industrious, or productive than rival workers present on the job. Sometimes employment decisions are actually made “on the merits.” Sometimes managers’ decisions are uncontaminated by illegitimate motives. And, as an extensive body of “mixed motive” jurisprudence recognizes, even decisions that are partly influenced by bias, chance, or arbitrary considerations might still be justified by other factors that are rationally related to personnel management.⁸⁸

In the world of employment discrimination law, these possibilities are assumed to exist, and they matter. The doctrines surrounding the law of discrimination are designed to distinguish legitimate and/or nondiscriminatory outcomes from those that are illegitimate and/or discriminatory—an often difficult determination that is notoriously prone to error.⁸⁹ However, this distinction—which is central to anti-discrimination law—receives insufficient attention by scholars of unconscious bias who believe that discrimination is pervasive.⁹⁰ The Kang and Banaji article is emblematic. The authors disparage the very notion of merits-based outcomes, implying that real-world determinations are rarely uncontaminated by sexual or racial identity. It thus follows that virtually all workplace practice is presumptively unlawful. They state:

Although perceivers assume that their judgments are based “on the merits”—in other words on the basis of qualities that the target in fact exhibits—the truth is more complicated. Even if we lack animus, intention to discriminate, or self-awareness of bias, our judgments of others may still lack “mental due process.” On subjective measures of merit, the perceiver’s (evaluator’s) expectations guide what she actually sees in the target In more plain language, if we expect someone to be violent, we will likely see violence when presented with ambiguous behavior.⁹¹

88. See generally Robert Belton, *Mixed-Motive Cases in Employment Discrimination Law Revisited: A Brief Updated View of the Swamp*, 51 MERCER L. REV. 651 (2000).

89. See generally Richard Goldstein, *Two Types of Statistical Errors in Employment Discrimination Cases*, 28 JURIMETRICS J. 32 (1985).

90. See, e.g., Ralph Richard Banks & Richard Thompson Ford, *(How) Does Unconscious Bias Matter?: Law, Politics, and Racial Inequality*, 58 EMORY L.J. 1053, 1100–01 (2009) (noting that, for unconscious bias theorists who believe “racial discrimination remains widespread,” the existence of “[racially] skewed outcomes alone become potentially persuasive evidence of discrimination”).

91. Kang & Banaji, *supra* note 72, at 1085 (footnote omitted).

This passage is notable more for what it fails to say than for what it does. People think they are objective, but their judgments “may lack ‘mental due process.’” Their perceptions are “likely” to be skewed by expectations, rather than to reflect an objective reality. The authors stop short of asserting categorically that impartial assessments are impossible, and they never say outright that merits-based judgments do not exist. But the clear implication is that, in most contexts that matter, the notion of a “judgment on the merits” is a delusion and a snare, or at best an ideal state that is rarely, if ever, achieved. In short, although the authors never say “never,” they disparage the idea that merit routinely governs decisions in the real world. Beyond that their exposition is singularly unhelpful. They fail to provide (1) a quantitative assessment of the degree to which decisions are distorted; (2) a benchmark for ascertaining when bias is operating and when it is not; or (3) a protocol for applying these observations to the concrete task of deciding when laws against discrimination have been violated and who is entitled to relief.

Kang and Banaji are not the only scholars to slight the quest for objective evidence of discrimination in favor of extravagant extrapolations from bias research. Another example is a recent article bemoaning the low success rate of anti-discrimination lawsuits.⁹² It is well known that litigants alleging discrimination prevail at lower rates than plaintiffs in most other areas of civil litigation. Katie Eyer attributes this pattern to people’s unwillingness to believe that social actors have violated anti-discrimination laws, which translates into fewer jury verdicts for plaintiffs.⁹³ She bases this conclusion on research that allegedly shows a reluctance to “see” discrimination in a “wide array of factual circumstances—ranging from traditional disparate treatment to more complex forms of bias.”⁹⁴ To support this assertion, Eyer relies on evidence she claims demonstrates that “most individuals think of discrimination as a phenomenon that is explicit, restricted in its manifestations, and generally unlikely to occur.”⁹⁵ This means that “in all but the most compelling factual circumstances, most people believe that some measure of merit—such as effort or ability—is a more likely explanation for why minorities fail than the possibility of discrimination.”⁹⁶ In short, Americans share a fundamental belief in meritocracy and think that its principles dominate key spheres of social and economic life. These “prevalent background beliefs,” she maintains, “account—at least in part—for the dismal odds that discrimination litigants face in litigating actual cases in the courts.”⁹⁷

Eyer’s article, however, contains almost no discussion of the precise legal standards that discrimination plaintiffs must meet or the types of evidence they must adduce to support allegations of unlawful disparate treatment. Moreover, the author’s argument is oblivious to the need for objective criteria by which the merits of

92. Katie R. Eyer, *That’s Not Discrimination: American Beliefs and the Limits of Anti-Discrimination Law*, 96 MINN. L. REV. (forthcoming 2011), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1780235.

93. *Id.* (manuscript at 3–4).

94. *Id.* (manuscript at 4).

95. *Id.*

96. *Id.*

97. *Id.*

discrimination claims are to be assessed. These are serious shortcomings. It should be obvious that, without some independent measure of how much discrimination actually takes place, it is impossible to know whether plaintiffs in discrimination suits prevail too seldom or too often. In short, Eyer provides no method for ascertaining how often plaintiffs *should* prevail. According to Eyer, however, there is no need for an objective baseline or any empirical reality check. She just assumes there is more discrimination out there than juries are willing to find.

Without objective validation, Eyer's argument is fatally weak. Although plaintiffs' success rate in discrimination claims is low relative to other types of suits, there is no *a priori* reason to think that outcomes will match across areas of litigation, and there may be good reasons why they do not. Why might large numbers of discrimination suits lack merit? Persons who are deprived of a desired and desirable condition of employment—such as a sought-after job or promotion—will often feel aggrieved. A person's fate on the job affects his livelihood, future, identity and self-worth. Attributing negative outcomes at work to discrimination rather than to one's own shortcomings—or just to bad luck—spares the ego and softens the psychological blow of disappointment. Moreover, a high error rate in plaintiffs' attributions of discrimination could well be due to the degree of information asymmetry that is endemic to the workplace. Employees are rarely in a position objectively to compare themselves to fellow workers or job candidates and are seldom privy to all the considerations behind staffing decisions. In sum, there are many reasons why employees might "see" discrimination when there is none. In the absence of systematic data, the existence and magnitude of these influences are, of course, necessarily speculative. But the notion that employees over-interpret the facts in their own favor is no less plausible than Eyer's suggestion that jurors and judges refuse to acknowledge discrimination where it exists. In the end, Eyer's assertion that the number of successful discrimination suits is "too low"—as opposed to "too high" or "about right"—evinces a belief that she shares with other scholars of implicit bias. If such bias is ubiquitous and the resulting discrimination pervasive, why bother to actually prove it? Why go through the arduous process of demonstrating that discrimination is the cause of any particular unfavorable decision? Why take on the burden of excluding rival explanations? For true believers in the pervasive influence of unconscious bias, there is no need to engage these onerous tasks.

B. *Remembering the Supply Side*

The cavalier attitude towards baselines and the obliviousness to objective criteria that are on display in the unconscious bias scholarship are disturbing. As noted, adverse outcomes for individuals or for groups in the employment context can have multiple possible explanations. The claim that stereotypical thinking routinely contaminates personnel decisions must be evaluated empirically. The key complication that bedevils this area is that of the "supply side"—the well-documented disparities between groups in American society in skills, human capital development, qualifications, and behavior.⁹⁸ In the context of discrimination law and doctrine, such differences are

98. See generally Wax, *supra* note 12.

routinely ignored or assumed away. Indeed, a pervasive assumption in discrimination law is that groups will be represented in most jobs in proportion to their share of the population.⁹⁹ Absent illicit conduct, workplace demographics should reflect the background population, and any deviation is suspect. It follows that statistical imbalances must be due to discrimination, rather than some other factor. Likewise, unfavorable decisions for individual minorities are presumptively illegitimate.

These understandings inform both disparate treatment and disparate impact doctrines in employment discrimination law. As I have argued elsewhere,¹⁰⁰ they lie at the heart of the disparate impact theory under Title VII of the Civil Rights Act. As first articulated in *Griggs v. Duke Power*,¹⁰¹ and carried forward in subsequent cases and EEOC regulations,¹⁰² the assumption of equal workplace representation finds expression in the so-called “4/5 rule.” This rule dictates that a business is presumptively liable for discrimination if the percentage of minority employees is less than 80% of candidates hired from the majority white population.¹⁰³ Yet if the minority working-age population—or minority candidates for a particular job—are on average less skilled or qualified than other segments of the workforce, businesses that engage in competitive meritocratic staffing will routinely violate the 4/5 rule.¹⁰⁴ Although employers can defend against liability by showing that their procedures are “job related,” meeting this standard is notoriously difficult, expensive, and fraught with uncertainty.

Similar observations hold for disparate treatment allegations. At the point of hiring, the decision to offer a job to a non-minority candidate rather than, for example, a black candidate would only be suspect if the candidates were “similarly qualified.” However, the selection criteria that employers often use—including years of education, prior experience, and the results of screening procedures such as interviews—are imperfect proxies for ability and only partial predictors of future performance. Years of education completed, or the possession of a high school or college diploma, fail to distinguish between persons who have developed different degrees of proficiency or who have learned more or less in school. In fact, there is ample evidence that blacks on average know less and acquire fewer skills than whites with similar years of education.¹⁰⁵ And there is also good evidence that academic achievement and tests of

99. See *Int'l Bhd. of Teamsters v. United States*, 431 U.S. 324, 340 n.20 (1977) (“[A]bsent explanation, it is ordinarily to be expected that nondiscriminatory hiring practices will in time result in a work force more or less representative of the racial and ethnic composition of the population in the community from which employees are hired.”).

100. See Amy L. Wax, *Disparate Impact Realism*, 53 WM. & MARY L. REV. (forthcoming 2011).

101. 401 U.S. 424 (1971).

102. Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607.4(D) (2010).

103. Wax, *supra* note 100, (manuscript at 2).

104. *Id.* (manuscript at 4–5).

105. See, e.g., Eric A. Hanushek, *How Well Do We Understand Achievement Gaps?*, FOCUS, Winter 2010, at 5, 6–7 (citing data from national tests showing “stunning” gaps in academic performance between black and white high school students, with the average seventeen-year-old black student “achieving at the 22nd percentile” for math and reading); Sharon Otterman, *37 Percent of New York Graduates Are College-Ready, Data Show*, N.Y. TIMES, June 14, 2011, at A23 (discussing New York State study on math and English scores which found that 51% of white students and 56% of Asian-American students were “college ready” after completing high school, versus only 13% of black students and 15% of Hispanic students); see also

cognitive skill are fairly good predictors of workplace productivity and success.¹⁰⁶ Given this data, the job performance of blacks and whites matched on years of education will tend to diverge, and thus the two groups will not—and should not—be hired at the same rates. Likewise, blacks and whites will on average fail to achieve the same average level of success on the job, which will likely affect pay and promotions.¹⁰⁷ Of course, these general observations reveal nothing about individual cases or whether a particular employee has been treated justly or is the victim of discrimination. But they do justify wariness in inferring discrimination from disparate outcomes, and they should encourage scholars to take alternative explanations for existing patterns seriously.

V. CURING UNCONSCIOUS BIAS: THE GAP BETWEEN PROPOSED REFORMS AND THE EVIDENCE THEY CAN REDUCE DISCRIMINATION

As legal scholarship on unconscious bias has mushroomed, the question of how to address inadvertent bias has come to the fore. Like the scholarship touting the ubiquity of implicit stereotyping, the literature putting forth solutions¹⁰⁸ has slighted the distinction between speculation and evidence. Once again, the obliviousness to baselines and the virtually complete absence of attention to supply-side factors is striking. Kang and Banaji, for example, propose a general set of recommendations for “proactive structural interventions,” on a “collective public health” model, that bypasses reliance “solely on potential individual litigation.”¹⁰⁹ Drawing on the “social

RICHARD ARUM & JOSIPA ROKSA, *ACADEMICALLY ADRIFT: LIMITED LEARNING ON COLLEGE CAMPUSES* 111–112 (2011) (noting that black students on average learned less during college than whites even when entering academic credentials were controlled); THOMAS J. ESPENSHADE & ALEXANDRIA WALTON RADFORD, *NO LONGER SEPARATE, NOT YET EQUAL: RACE AND CLASS IN ELITE COLLEGE ADMISSIONS AND CAMPUS LIFE* 226–29 (2009) (discussing research on group differences in scores and academic achievement); Richard C. Hunter & RoSusan Bartee, *The Achievement Gap: Issues of Competition, Class, and Race*, 35 *EDUC. & URB. SOC'Y* 151, 154–56 (2003) (same).

106. See Frank L. Schmidt, *The Role of General Cognitive Ability and Job Performance: Why There Cannot Be a Debate*, 15 *HUMAN PERFORMANCE* 187, 187 (2002) (noting “overwhelming research evidence showing a strong link between general cognitive ability . . . and job performance”); see also Philip L. Roth & Philip Bobko, *College Grade Point Average as a Potential Selection Device: Ethnic Group Differences and Potential Adverse Impact*, *J. APPLIED PSYCHOL.* 399, 399 (2000) (discussing correlation of grade-point average to job performance); Paul R. Sackett et al., *High-Stakes Testing in Higher Education and Employment*, 63 *AM. PSYCHOLOGIST* 215, 216–19 (2008) (discussing predictive power of admission test results); Wax, *supra* note 12, at 994 (citing Mark Kelman, *Concepts of Discrimination in “General Ability” Job Testing*, 104 *HARV. L. REV.* 1157 (1991)) (discussing pros and cons of general aptitude job testing as method for predicting job performance).

107. See Paul R. Sackett & Steffanie L. Wilk, *Within-Group Norming and Other Forms of Score Adjustments in Preemployment Testing*, 49 *AM. PSYCHOLOGIST* 929, 934 (1994) (discussing studies in organizational psychology reporting that blacks receive on-the-job performance ratings that average about .3-.4 standard deviations below ratings for white workers).

108. For examples of this literature, see Tristin K. Green, *Race and Sex in Organizing Work: “Diversity,” Discrimination, and Integration*, 59 *EMORY L. REV.* 585 (2010); Tristin K. Green & Alexandra Kalev, *Discrimination-Reducing Measures at the Relational Level*, 59 *HASTINGS L.J.* 1435 (2008); Christine Jolls & Cass R. Sunstein, *The Law of Implicit Bias*, 94 *CAL. L. REV.* 969 (2006); Kang & Banaji, *supra* note 72.

109. Kang & Banaji, *supra* note 72, at 1080.

contact hypothesis”—the theory that integration reduces stereotyping, decreases automatic biases, and promotes equality norms¹¹⁰—Kang and Banaji advocate the widespread adoption of affirmative action as a “tie breaker” and a mechanism for achieving more diversity and social mixing.¹¹¹

These recommendations have serious drawbacks. First, it is far from clear whether affirmative action will have much effect on the incidence of unconscious bias at all, let alone whether it will reduce harms from discrimination, as the authors predict. The authors offer no studies that actually document such outcomes. Nor do they show that the people who will benefit from affirmative action are those who would otherwise fall victim to unconscious ill treatment. Likewise, they fail to demonstrate that affirmative action would rectify any discrimination to the appropriate degree, as opposed to conferring advantages that are remedially unjustified. In short, the authors assume everything and demonstrate nothing. Second, the evidence indicates that using affirmative action as a tie-breaker will not measurably alter the demographic composition of the workplace, academic institutions, or other competitive social contexts. As noted, academic and workplace imbalances are largely traceable to persistent and significant disparities by race in skill and developed human capital.¹¹² The problem is that there are relatively fewer qualified blacks in the population than persons from other demographic groups. Blacks are thus seldom “tied” with white or Asian candidates for competitive positions.

One well-documented example is law school admissions. Blacks significantly underperform whites on the Law School Admission Test (LSAT), which is an important criterion for admission to law schools nationwide. And very few blacks achieve scores (about 165 and above) that would earn admission at elite institutions.¹¹³ Another example is recruitment into the military. The U.S. government has long used entrance exams that measure or draw on cognitive ability to determine admission to the armed forces and assignments to desirable positions within it. These screens have a pronounced disparate impact by race. According to a recent report, while 16% of otherwise qualified whites applying to enter the military (i.e., men and women with a high school degree and no serious criminal record) failed to achieve the required minimum score on the Armed Forces Qualifying Test (AFQT), 39% of black applicants scored below the cutoff.¹¹⁴ Among candidates admitted, over 43% of white inductees, but fewer than 18% of black inductees, scored high enough (in the top two categories) to qualify for special technical training and placement in elite service jobs.¹¹⁵ It should

110. For a discussion of the social contact hypothesis, see Jerry W. Robinson, Jr. & James Preston, *Equal Status Contact and Modification of Racial Prejudice: A Reexamination of the Contact Hypothesis*, 54 *SOCIAL FORCES* 911 (1976).

111. Kang & Banaji, *supra* note 72, at 1100–02.

112. See *supra* notes 105–07 and accompanying text.

113. See, e.g., Daria Roithmayr, *Tacking Left: A Racial Critique of Grutter*, 21 *CONST. COMMENT.* 191, 192 (2004) (noting median black LSAT score in 2002–2003 of 142.2 vs. 153.9 for whites, and observing that blacks make up fewer than one percent of candidates scoring above 165 on test).

114. CHRISTINA THEOKAS, EDUC. TRUST, SHUT OUT OF THE MILITARY: TODAY’S HIGH SCHOOL EDUCATION DOESN’T MEAN YOU’RE READY FOR TODAY’S ARMY 3, 5 (2010), available at http://mldc.whs.mil/download/documents/Readings/ASVAB_4.pdf.

115. *Id.* at 5.

be obvious that using affirmative action as a “tie breaker” will do little to alter these stark disparities. This observation is important, because service in the military functions as a training ground and stepping-stone into the workforce for many people without a college degree. Yet blacks are disproportionately excluded because they have not achieved the minimum competence necessary to function in the military setting. In sum, lack of preparation, rather than the failure to prefer equally qualified blacks, is responsible for most of the underrepresentation of blacks in desirable positions in government, business, and the professions. “Tie breaking” affirmative action will do little to change this. Much more drastic affirmative action would be needed to achieve a measurable effect. Indeed, the most important implication of the underlying data is that unconscious biases—and discrimination generally—are not the primary reason for blacks’ poor outcomes or relative lack of success. Rather, large and well-documented skill differences more than account for observed patterns of group representation in the workplace and society generally. Without better quantitative evidence on the role of implicit biases relative to other measurable “supply side” factors, there is no reason to believe that banishing or alleviating unconscious bias will have much effect on racial gaps in the real world overall.

Finally, although Kang and Banaji (as well as other scholars) rely heavily on the “social contact” hypothesis in formulating approaches to unconscious bias,¹¹⁶ the evidence on the benefits of interracial mixing is decidedly mixed. Even Kang and Banaji admit that strategies to increase diversity and promote group interaction are unpredictable and potentially counterproductive. They note the possibility that affirmative action may “strengthen[] identification with and resentment across race[s].”¹¹⁷ They cite evidence that “intergroup contact may not counteract negative stereotypes,” but might in some cases promote them.¹¹⁸ Given the unsettled nature of the research, and the paucity of data on how interracial mixing operates in particular settings, an aggressive stance towards mandated integration as a cure for unconscious bias is decidedly premature.

Two other recent articles also exemplify the trend towards sweeping recommendations based on insufficient evidence. As with others who have written in this area, Tristin Green and Alexandra Kalev bemoan the limitations of anti-discrimination law and the restricted potential of conventional litigation to address the types of unconscious bias and stereotyping that are claimed to infect the contemporary workplace.¹¹⁹ According to Green and Kalev, employers should be encouraged or

116. Kang & Banaji, *supra* note 72, at 1100–08. For a discussion of the social contact hypothesis, see generally Robinson & Preston, *supra* note 110.

117. Kang & Banaji, *supra* note 72, at 1101.

118. *Id.* at 1104–05; see also Katharine Bartlett, *Making Good on Good Intentions: The Critical Role of Motivation in Reducing Implicit Workplace Discrimination*, 95 VA. L. REV. 1893, 1951 (2009) (noting that contact with people from other groups “can either reinforce bias or reduce it”).

119. See generally Green & Kalev, *supra* note 108. The authors draw heavily on a number of influential articles in the field, including Susan Sturm, *Second Generation Employment Discrimination: A Structural Approach*, 101 COLUM. L. REV. 458 (2001) and Linda Hamilton Krieger, *The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity*, 47 STAN. L. REV. 1161 (1995); see also Suzanne B. Goldberg, *Discrimination by Comparison*, 120 YALE L.J. 728 (2011); Tristin

required to take informal steps designed to reduce unconscious discrimination at the relational level. Legislatures should even consider imposing liability for “failing to address relational sources of discrimination.”¹²⁰ The “debiasing” measures they recommend include affirmative action to create a critical mass of “role models” and “countertypical exemplars,” diversity training, formalization and transparency in decisionmaking, an increase in cooperative teamwork, and a reduction in hierarchy and status differentials on the job.¹²¹ These recommendations are too sprawling to analyze in detail here, but suffice it to say that they suffer from breathtaking scope and maddening vagueness. They also rest on say-so. The authors advance little or no evidence that the recommended measures will achieve their stated purpose and provide no criteria for assessing whether or not they do.

Katharine Bartlett’s recent foray into the “debiasing” project is more forthright in acknowledging the limitations of the data.¹²² She concedes that the systematic structural reforms proposed as an antidote to implicit bias have never actually been proven effective.¹²³ She also accepts and defends the fundamental paradigm of disparate treatment law, which demands that discrimination be demonstrated through a causal relationship between impermissible factors and adverse outcomes.¹²⁴ Along those lines, Bartlett, criticizes many proposals for “short-circuit[ing] basic issues of proof and causation and second-guess[ing] a wide range of employer decisions.”¹²⁵ Accordingly, she faults proposed reforms that impose employer liability “based on potential rather than actual discrimination”; that “abandon the necessity for establishing a causal link between an employment decision and a victim’s race, sex or other protected characteristic”;¹²⁶ and that are “designed to hold employers liable for workplace procedures and structures that could give rein to unconscious bias, even if not shown to do so.”¹²⁷ In short, she deplores the suggestion that discrimination should be found by “a rule that presumes it.”¹²⁸

Bartlett also asserts that some of the “structural” reforms that have been proposed by scholars of unconscious bias can potentially have a “negative effect[] . . . on the very discrimination they seek to reduce.”¹²⁹ Specifically, interference with a range of management prerogatives can engender resentment and a sense of unfairness within the workplace, thus “undermining the conditions necessary to motivate people to want to avoid implicit discrimination.”¹³⁰ She singles out diversity training and diversity evaluations as ineffective and as patronizing workers by implying they “need to be

Green, *A Structural Approach as Antidiscrimination Mandate: Locating Employer Wrong*, 60 VAND. L. REV. 849 (2007).

120. Green & Kalev, *supra* note 108, at 1457 & n.140.

121. *Id.* at 1438–54.

122. Bartlett, *supra* note 118, at 1967.

123. *Id.* at 1960–72 (critiquing reform proposals).

124. *Id.* at 1956–1960.

125. *Id.* at 1930.

126. *Id.* at 1957.

127. *Id.* at 1958.

128. *Id.*

129. *Id.* at 1930.

130. *Id.*

taught not to discriminate because right now they do not know any better.”¹³¹ She is critical of affirmative action because “measures that appear to favor members of protected groups at the expense of identifiable, seemingly deserving majority workers . . . tend to provoke resistance.”¹³²

Despite these criticisms, Bartlett looks upon some “workplace culture” changes with favor. She claims that workplace collaboration, the reduction of “status differentials,” and “job rotation programs” can reduce stereotyping and improve prospects for women and racial minorities in the workplace.¹³³ That is because they “focus on promoting people’s good intentions,” and on enhancing “contact and cooperative attitudes” among people from different groups.¹³⁴

It is hard to know how to evaluate Bartlett’s selective rejection and embrace of different “debiasing” proposals. Despite her relative circumspection, she endorses reforms (team building and collaborative cooperation) for which the evidence is as thin or equivocal as for the reforms she denigrates. The research she cites provides no objective measure of actual discrimination or bias, nor do the studies she relies upon show its abatement—they simply claim to register more positive attitudes or document more favorable actions towards minorities without demonstrating that these are objectively justified by neutral, valid, non-race-based criteria.¹³⁵ At bottom, the efficacy of Bartlett’s proposed measures (and others) remains a matter of speculation. Specifically, whether particular interventions will actually produce less unlawful discrimination—that is, fewer instances of injurious behavior “because of” race or sex in any given context—remains to be seen. It is ironic that Bartlett, who criticizes some proposals for dispensing with proof of actual discriminatory causation, fails to discharge that burden herself. To be sure, her recommendations are informal: she does not propose requiring structural reform as a matter of law. But she nonetheless asks us to believe that the changes she advocates will have the predicted effects. Bartlett cannot escape the fact that there is no reliable gauge of stereotyping and bias in the workplace, and no objective baseline for determining when those influences are operating. Creating more diversity and putting more minorities in positions of power—although perhaps desirable in itself—does not establish that discrimination and bias have abated or disappeared. Likewise, less diversity does not mean that bias is operating.

VI. UNCONSCIOUS BIAS: WHAT IS TO BE DONE?

A more productive approach might be to step far back from the debate and think about ultimate objectives. What are the goals of anti-discrimination law and workplace reform more generally? In this regard, race and gender present somewhat different issues, if only because the potential causes—and cures—for observed race and gender disparities are different.

131. *Id.* at 1960, 1961.

132. *Id.* at 1968.

133. *Id.* at 1960–65.

134. *Id.* at 1960–63.

135. *See id.* at 1960–71 (discussing the evidentiary basis for Bartlett’s proposed reforms).

In the case of race, inequalities are widespread across multiple domains. Ralph Banks and Richard Ford have recently confronted this fact in the context of reviewing the large body of work on unconscious bias.¹³⁶ Their discussion comes close to achieving the tentative and measured response that the current state of the art warrants. First, the authors make some insightful points about the “ascendance of unconscious bias” as a dominant theme in race-discrimination scholarship.¹³⁷ They speculate that the growing focus on inadvertent rather than deliberate bias can be ascribed to the political appeal of identifying unconscious processes as the dominant source of racial prejudice and discrimination in society today.¹³⁸ People are unaware of their inadvertent stereotypes and have little control over them, so they appear less blameworthy for possessing and acting on them. As Banks and Ford explain, “[b]ecause the unconscious bias discourse does not imply that one is a liar or a racist, there is less reason to resist the claim that implicit bias thrives somewhere within one’s self.”¹³⁹ Thus “people may be willing to acknowledge the possibility of unconscious bias within themselves even as they would vigorously deny harboring conscious bias.”¹⁴⁰

On the other hand, Banks and Ford note that there are costs to assigning a central role to unconscious, as opposed to animus-driven, racism. The lesser blameworthiness of unconscious motives can diminish people’s sense of responsibility and thereby undermine the imperative to address and rectify racism.¹⁴¹ The authors also fear that focusing on unconscious bias will intensify efforts to exert intrusive and undesirable forms of social control. They identify as major risks of “unconscious bias discourse” a trend towards “technocratic authoritarianism” and the increased use of heavy-handed efforts to “enforce a norm of technically rational decision making.”¹⁴² The authors further observe that “unconscious bias theory makes racism seem more a medical problem than a social one.”¹⁴³ This invites systemic and coercive interventions analogous to public health measures, which can seriously compromise personal freedom and social flexibility.

Finally, Banks and Ford are wary of the discrepancy between claims of unconscious bias and the evidence of resulting discrimination. In line with the concerns expressed above, the authors identify “the pivotal empirical question” as “whether [unconscious reactions] prompt discriminatory behavior.”¹⁴⁴ In particular, they ask whether the IAT is a good predictor or measure of propensity to discriminate.¹⁴⁵ These authors candidly admit that, at the present juncture, the answer to this question must be

136. See generally Banks & Ford, *supra* note 90.

137. *Id.* at 1054.

138. *Id.* at 1104 (noting that the accusation of widespread unconscious bias is “more politically palatable than parallel claims” about pervasive animus-driven bias).

139. *Id.*

140. *Id.* at 1054.

141. *Id.* at 1103–06.

142. *Id.* at 1116.

143. *Id.* at 1118.

144. *Id.* at 1111.

145. *Id.*

a resounding no.¹⁴⁶ They assert that the “ambiguities and uncertainties” in the science make IAT evidence, which consists mainly of laboratory observations, a poor guide to behavior in the real world.¹⁴⁷ They also note that this weakness has been “glossed over in most popular accounts of the IAT.”¹⁴⁸ The authors conclude that, although the IAT “may eventually prove a useful predictor of discriminatory behavior, [it] cannot now be comfortably described as such.”¹⁴⁹ It follows that tests designed to gauge unconscious bias cannot accurately identify perpetrators of discrimination. Nor can they tell us whose decisions are and will be tainted by impermissible considerations of race. Indeed, there is no indication that laboratory measures of inadvertent bias provide any information that might be useful or dispositive in legal claims of discrimination.

In the end, Banks and Ford argue that the obsession with inadvertent stereotyping is a distraction from the real causes—and cures—for racial inequality. Although not denying “the existence of unconscious bias,” the authors “do doubt that racial bias explains all or even most of the racial injustices that plague our society.”¹⁵⁰ In their view, “racial injustice . . . is a matter of pervasive substantive inequalities, the amelioration of which should be the goal of reform efforts.”¹⁵¹ For them, the needed improvements are unlikely to come from searching for instances of unconscious bias or from working to establish its influence. The authors recognize that “the embrace of the unconscious bias idea” might be viewed by some scholars as a way to generate “support for policies that would further the goal of substantive equality.”¹⁵² But they do not believe this strategy will ultimately prove effective, and they are sensitive to its considerable costs. They instead recommend that the goals of “integration, reparations, or equal outcomes” be pursued directly.¹⁵³

The Banks and Ford article is useful in reviewing the shortcomings of unconscious bias research, in explaining why skepticism towards claims about unconscious discrimination is justified, and in expressing appropriate concern that the obsession with unconscious bias is misplaced. Beyond that, however, the authors are vague on the particular inequalities they want addressed and the measures that should be taken to correct them. Without a close analysis of those inequalities (economic, personal, familial, educational, occupational, etc.) and the strategies proposed to tackle them, the authors' recommendations cannot adequately be assessed. Some general observations are in order, however. Banks and Ford repeatedly use the word “injustice” to describe the inequalities and adverse social conditions that plague minorities. But they fail to grapple with the possibility that not all group inequalities are necessarily “unjust.” In other words, not all disparities—including, possibly, those that emerge between different groups—can be ascribed to unfair social conditions, as opposed to differences in effort, restraint, preferences, performance, or bad luck. Accordingly, not

146. *Id.*

147. *Id.*

148. *Id.*

149. *Id.*

150. *Id.* at 1113.

151. *Id.* at 1110.

152. *Id.* at 1113.

153. *Id.*

all will warrant correction on grounds of justice. In addition, in issuing an urgent call for a social response to racial inequalities, Banks and Ford fail to acknowledge the possibility that government action and social programs may prove ineffective against many observed group differentials, or may even do more harm than good.¹⁵⁴ Finally, their exhortation to drastic action raises the very concerns—of overbearing government and pervasive social control—that they deplore in responses to unconscious bias.

VII. PERSISTENT INEQUALITIES: REMEMBER THE SUPPLY SIDE

Banks' and Ford's concern with eradicating racial inequalities brings us back full circle to the problem of the supply side. What are the causes of group differences in outcomes, and how can they be effectively addressed? Legal scholarship looks first to discrimination, but a clear-eyed view of the entire picture reveals that disparities by race and class have more complex roots. This Article has focused primarily on racial inequality, where the supply side looms large. The relative underdevelopment of human capital in the black community, for instance, could help explain persistent inequalities in the workplace, and might account for a large portion of existing gaps.¹⁵⁵ Ford and Banks do not deny this. Rather, they appear to see it as a problem to be solved, but say very little about how to do that. Their acknowledgment of the influence of the supply side, however indirect, returns us to the pivotal issue in anti-discrimination law—which is causation. The numbers are out of whack, and the outcomes are disparate. There may be discrimination, but there may also be other factors at work. The relative importance of each is critical, and sorting this out is a matter of determining causation. The challenge is daunting, but the problem cannot be avoided or waved away. Without establishing that race (or sex, or some other forbidden factor) has entered into a decision, there can be no credible accusation of discrimination.

In resolving the allegations in a case like *Wal-Mart*, a court would have to decide whether a particular worker was the victim of discrimination or whether she was terminated, denied a raise, or passed over for a promotion for some other reason. "Structural problems" may indeed plague the *Wal-Mart* workplace. Perhaps, as alleged, the company's procedures render some personnel decisions vulnerable to bias. But that observation tells us nothing about how many—and which—of the myriad women who have worked for Wal-Mart have been unlawfully discriminated against, and how many have not.

The answers to these questions are likely to prove elusive, because excluding alternative explanations for adverse outcomes in the workplace or other social spheres is always difficult and sometimes impossible. At this juncture, the limitations of anti-discrimination law are manifest. Indeed, many commentators have noted the narrow scope and methodological strictures inherent in the legal doctrine.¹⁵⁶ Perhaps the

154. See generally AMY L. WAX, RACE, WRONGS, AND REMEDIES: GROUP JUSTICE IN THE 21ST CENTURY (2009) (arguing that many racial inequalities today elude policy-based correction).

155. See *supra* notes 105–07 and accompanying text for a discussion of the academic achievement gap between blacks and whites and the implications for job performance.

156. E.g., Goldberg, *supra* note 119; Krieger, *supra* note 119; Selmi, *supra* note 34; Sturm, *supra* note 119.

requirements are too stringent and should be relaxed. But it is very hard to translate that recommendation into specific, well-defined reforms that serve their stated purpose—which is to distinguish legitimate decisions from illegitimate ones. It is one thing to find fault with the current system, but quite another to devise an alternative that is fair to all concerned. Far-reaching transformations of the workplace—including the wholesale reorganization of work and sweeping shifts in how employees interact and are evaluated—are now the recommendations of choice.¹⁵⁷ Whether such untested and potentially costly measures will be implemented or will prove worthwhile is as yet unknown. But if Ford and Banks (and this author) are correct—that is, if unconscious bias and discrimination are relatively minor sources of group inequality today—the enduring solutions will surely lie elsewhere.

157. E.g., Kate Bartlett, *supra* note 118, at 1926–30 (reviewing proposals); Tristin K. Green, *The Future of Systemic Disparate Treatment Law*, 32 BERKELEY J. EMP. & LAB. L. (forthcoming 2011) (manuscript at 5, 36), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1793425 (recommending a “context model” that imposes liability for policies that are “likely to be producing” disparate treatment).