Institute for Law & Economics
University of Pennsylvania Law School
Research Paper No. 08-10

Public Law and Legal Theory
University of Pennsylvania Law School
Research Paper No. 08-11

The Discriminating Mind: Define It, Prove It

Amy L. Wax
University of Pennsylvania Law School

CONNECTICUT
LAW REVIEW

VOLUME 40 MAY 2008 NUMBER 4

Article

The Discriminating Mind: Define It, Prove It

AMY L. WAX

Differential group achievements in competitive spheres like business, government, and academia, in conjunction with professed organizational commitments to fairness and equal opportunity, fuel claims that unconscious discrimination operates widely in society today. But attempts to blame disparities by race or sex on inadvertent bias must be approached with caution in the current climate. Many allegations concerning unconscious discrimination do not properly allege category-based treatment at all but rather target the disparate impact, or differential effects, of category-neutral criteria. Such impacts often reflect well-documented "supply side" disparities between groups in human capital development, qualifications, and behavior. These patterns are not most effectively addressed by focusing on unconscious processes, but rather by scrutinizing neutral practices for efficiency and social usefulness and also by attempting to eliminate underlying group differences in the ability to compete for social rewards.

Likewise, allegations of unconsciously motivated disparate treatment, which are based on the contention that race or sex plays a causal role in social outcomes, should be scrutinized for alternative, non-discriminatory explanations for observed disparities, including "supply side" differences between groups. In addition, some disparities attributed to unconscious bias could just as well be explained by old-fashioned "statistical" or "rational" discrimination, which is also fueled by real, average, observable differences in performance by race or sex. In general, sweeping and categorical claims of unconscious discrimination are unwarranted without specific evidence that this process is actually operating in a given case. Such evidence is hard to come by. In many cases, supporting such claims requires excluding alternative explanations—including "supply side" explanations—for observed disparities in group success.

ARTICLE CONTENTS

The Discriminating Mind: Define It, Prove It

AMY L. WAX*

I. INTRODUCTION

Like other participants in this conference in honor of Charles Lawrence's seminal Article,[1] I am deeply interested in the problem of inequality in our society—economic, racial, social and sexual. What are the sources of existing disparities, and how can and should we address them? These questions pre-occupy those assembled here today. But, as someone with a scientific background, I am wary of explanations that posit causal mechanisms for inequality and acutely aware of the difficulties of substantiating such claims. When the assertion is made—as it often is in discussions of inequality—that X causes Y, my first thought (and the thought, I hope, of any good social scientist) is: correlation is not causation. If X causes Y, where is the proof?

Of course the problems of proof and causation have been long-standing features of anti-discrimination law from the very beginning. But in the wake of the work of Charles Lawrence and his acolytes as well as the growing emphasis on the role of unconscious stereotyping and inadvertent discrimination in many areas of social life, these issues have taken on renewed urgency. The focus on unconscious bias has not only highlighted many uncertainties regarding issues of causation, correlation, and proof for discrimination claims, but has also renewed old terminological confusions. Some of the issues raised are longstanding, but others are new and peculiar to the conceptual difficulties surrounding claims of unconscious bias. Some of these have been addressed in my 1998 article, *Discrimination as Accident*,[2] and much of what I say there is still relevant to the conceptual and terminological conundrums posed by unconscious discrimination. On the issue of proof, however, there is much new water under the bridge, with an accumulation of empirical studies and new data, and renewed claims based on that evidence.

Before addressing the question of how unconscious discrimination claims—and discrimination claims generally—are substantiated (which is the main topic of this Article), I want briefly to consider—and I hope to

---

* Robert Mundheim Professor of Law, University of Pennsylvania Law School. I thank Jessica Weiss for excellent research assistance. All errors are mine.

[1] Charles R. Lawrence III, *The Id, The Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317 (1987).

[2] Amy L. Wax, *Discrimination as Accident*, 74 IND. L. J. 1129 (1999).

clarify—a few key conceptual and terminological issues that arise repeatedly in this context.

First, I begin with the well-known distinction between disparate treatment and disparate impact. Disparate treatment results when a personal attribute—like sex or race—is taken into account to affect decisions or to influence how individuals are treated by others. By definition, a person's racial or sexual identity is causally—and necessarily cognitively—implicated in the decision-making process. For disparate treatment to occur, the decision-maker must at some point observe, or at least be mentally exposed to, information about another person's identity. Without that input, there can be no disparate treatment. In cases of disparate impact, in contrast, race or sex does not causally enter into the decision-making process at all. Rather, application of neutral criteria for decision-making ends up having differential effects on different groups.[3]

The tried and true categories of disparate treatment and disparate impact, and the law that has come to apply to them, have not received universal endorsement. In particular, some have argued that covert, unconscious forms of discrimination fit uneasily within this scheme. The contention is that, because the old categories predate the rise of these new, subtle types of bias, present realities render them obsolete. My contention here is that this critique is unjustified and misguided. Properly understood, the categories of disparate treatment and disparate impact create a serviceable conceptual framework for analyzing all claims of bias in social life, whether deliberate or inadvertent, gross or subtle. Moreover, the old requirements for making out claims under these categories are still relevant and should continue to apply with full force to discrimination that is unconsciously motivated.

With respect to defining or identifying disparate treatment, I argued previously that there is no conceptual or theoretical reason to distinguish among decisions that are influenced by a person's racial identity on the basis of whether those decisions implicate conscious or unconscious mental processes or are the product of deliberate awareness or inadvertency.[4] If, for example, a person treats someone differently—and adversely—*because of* that person's race, then that would violate the plain

---

[3] With respect to distinguishing disparate treatment from disparate impact, there are borderline cases that are not easy to classify. One such case is that of neutral criteria that are selected because they disfavor one group. Although how to categorize such a case is subject to dispute, some commentators have argued for regarding this as a form of disparate treatment because race or sex actually enters into the decision-making process. On the role of intent in disparate treatment and disparate impact claims, see, for example, Michael Selmi, *Proving Intentional Discrimination: The Reality of Supreme Court Rhetoric*, 86 GEO. L.J. 279 (1997); David A. Strauss, *Discriminatory Intent and the Taming of* Brown, 56 U. CHI. L. REV. 935 (1989). While taking note of this possibility, the discussion herein will not centrally be concerned with this borderline case.

[4] *See* Wax, *supra* note 2, at 1137–38 (explaining that trait-based disparate treatment can still occur even when the actor is unaware of being influenced by the worker's race or sex).

terms of Title VII of the Civil Rights Act of 1964,[5] which forbids such conduct. That would constitute unlawful disparate treatment whether or not the actor is fully aware that another person's protected characteristic—his race, for example—has influenced the decision. Although, as a practical matter, an actor cannot engage in discrimination without having "observed" another person's race, that observation need not be conscious. There need be no element of awareness that race is a factor influencing conduct. All that is required to satisfy the plain terms of the statute is that race be causally implicated. This can happen consciously or unconsciously. Thus, the statute itself does not limit its application solely to *deliberate* treatment "because of" protected characteristics such as race.

This analysis also points to an important source of confusion surrounding the concept of "intent." The word "intent" in the context of discrimination law is both mischievous and misleading because it is ambiguous. It can be used to distinguish conscious from unconscious action. But it is also sometimes employed to differentiate action taken for a particular purpose from action taken despite that action's effect. Originally, the concept of intent was used in the law to express the second distinction—between claims of disparate treatment and those alleging only disparate impact. But it came to stand for the first distinction as well, leading to the erroneous conclusion that only *conscious* disparate treatment—that is, action *deliberately* based on a protected characteristic—is covered by the terms of anti-discrimination laws.

As noted, the concept of "intent" has frequently been employed to distinguish cases of disparate treatment—where the effect on racial groups was generally thought to be "intentional" or "intended"—and those of disparate impact, in which racial effects were a byproduct of other goals and thus "unintentional." The word "intent" thus made its appearance in cases such as *Washington v. Davis* and its progeny, where the Court's refusal to entertain constitutional disparate impact claims was expressed—infelicitously and ultimately confusingly—as the requirement that actionable discrimination be "intentional."[6] In light of this, it is a mistake to cite the equal protection cases for the proposition that actionable discrimination must be conscious or deliberate. In the wake of more precise insights about the possibility for unconscious—or "unintentional"—disparate treatment, and with careful attention to the core role of causation in anti-discrimination law, that interpretation is unjustified. Excluding unconscious discrimination from the category of potentially actionable conduct incorporates an overly restrictive notion of the causal element of disparate treatment.

---

[5] 42 U.S.C. § 2000e-2 (2000). That statute forbids certain types of decisions made because of race, sex, national origin, etc. *Id.*

[6] Washington v. Davis, 426 U.S. 229, 237–40 (1976).

Under the best reading of Title VII's terms, that statute covers unconscious disparate treatment. The equal protection guarantee should be read in consonance. For this reason, *Washington v. Davis* is best understood by leaving the concept of intent aside. That case stands for the proposition that the guarantee of equal protection only elicits heightened scrutiny for actions taken "because of" discrete protected characteristics. That is, it only protects against disparate treatment. Because disparate treatment because of race can be conscious or unconscious, such adverse treatment should be regarded as constitutionally suspect as well.[7]

The second confusion that has produced considerable debate concerns how to define racial bias or prejudice. Philip Tetlock and Hal R. Arkes have addressed this question in a widely discussed piece concerning the significance of the so-called Implicit Association Test (IAT).[8] As developed by Mahzarin Banaji and her colleagues at Harvard, the IAT measures the speed with which subjects associate negative words or concepts with images of people of different races. The association is thought to reveal unconscious processes similar to stereotyping or generalizations about groups. 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 of blacks.[9]

Tetlock and Arkes question whether the IAT is really a test of prejudice, bias or racism in the ordinary sense of those words. They point out that stronger and faster association of negative concepts with blacks will not necessarily spring from hostility. Rather, people might be aware of social stereotypes that they do not themselves endorse, or of demographic or social facts—related to black disadvantage—that they regret or believe society must correct.[10] Neither of these scenarios suggests the type of animus usually associated with racism as commonly understood—indeed, quite the contrary.

My view is that these observations, although valid and interesting, miss the point. The concern here should not be with the mental associations the IAT purports to measure. Nor does anything important

---

[7] Although conceding that the language of Title VII would permit imposing liability for unconscious disparate treatment, I previously argued against applying the statute to permit such claims largely on pragmatic grounds. The logic of my argument extends to claims of unconstitutional discrimination as well. Although I do explore the considerable practical difficulties of making out unconscious bias claims below, I do not again take up the normative question of the scope of anti-discrimination liability that was the focus of my previous work. *See* Wax, *supra* note 2, at 1226 (explaining the problems associated with detecting and remedying unconscious discrimination under anti-bias laws).

[8] Hal R. Arkes & Philip E. Tetlock, *Attributions of Implicit Prejudice, or "Would Jesse Jackson 'Fail' the Implicit Association Test?"*, 15 PSYCHOL. INQUIRY 257, 258–59 (2004).

[9] For more information on the IAT, see Anthony G. Greenwald & Linda Hamilton Krieger, *Implicit Bias: Scientific Foundations*, 94 CAL. L. REV. 945, 952–53 (2006).

[10] Arkes & Tetlock, *supra* note 8, at 258.

turn on resolving the semantic debate surrounding the meaning of "bias," "prejudice," or "racism." Rather, what really matters, and what ought to matter to law, is whether people are treated worse because of their race—or other protected characteristic, such as sex—in the real world. Specifically, the focus should not be on attitudes or sympathies, but traditionally on actionable discrimination. Once again, discrimination occurs when an individual is victimized by ill treatment that is causally linked to or based on a protected characteristic.

The Tetlock/Arkes critique of the IAT suggests a crucial related question: whether it is proper to equate unconsciously biased mental associations with the tendency to engage in unlawful discrimination.[11] Surely the answer to this question is "no." Biased thinking and attitudes, and mental processing of stimuli and concepts, are not the same as unlawful discrimination. It is important never to lose sight of this distinction. Racial discrimination is not about mental states. It is about social results and the causal basis for those results. Discrimination, whether conscious or unconscious, is abhorrent because it yields less favorable treatment of persons of one race in employment, education, public programs, political power, and other social arenas. Such adverse treatment affects status and well-being and arbitrarily exacerbates inequalities between groups.

On this view, what matters is not the mental state, whether conscious or unconscious. Mental states alone do not harm people. Adverse actions are what harm people. The key question is whether mental states generate discrimination.[12] It follows, I think, that we should only be concerned with unconscious bias—or conscious bias, for that matter—if it can be shown to produce real-world discrimination in the sense of disparate treatment as just defined.[13] Biased cognition thus commands our attention only if it reliably predicts or can be linked to actual discrimination.

But it is precisely in the matter of establishing that link that studies of unconscious bias, and the claims based on those studies in social science and legal scholarship, potentially get into trouble. That connection is too often assumed rather than demonstrated, and the hard work needed to show it—to actually prove it—is too often left undone. To the extent that efforts are made to tie bias to actual discrimination, they are often half-hearted.

---

[11] For such an opinion, see, for example, Jerry Kang, *Trojan Horses of Race*, 118 HARV. L. REV. 1489, 1514 (2005) (arguing the presence of implicit bias through IAT experiments and the proportional influence on real world behavior). *See generally* Kristin A. Lane et al., *Implicit Social Cognition and Law*, 3 ANN. REV. L. & SOC. SCI. 427 (2007).

[12] *C.f., e.g.*, R. Richard Banks et al., *Discrimination and Implicit Bias in a Racially Unequal Society*, 94 CAL. L. REV. 1169, 1186–87 (2006) (observing that "a mental state signifies bias if it consistently produces discrimination"). Although the concern with defining "bias" is a sideshow, the basic idea expressed here is a sound one: a mental state is of interest only if it actually produces, or is reliably associated with, discrimination in the real world.

[13] *See* discussion *infra* Part III.

The entire subject of proving unconscious discrimination is too often plagued with muzziness, imbued with wishful thinking, informed by political correctness, and oblivious to unpleasant facts.

In proving that unconscious mental categorization causes discrimination, the key is to isolate the effect of a personal characteristic—let's say race or sex—on the outcome of a social decision. It should go without saying that this necessarily requires excluding other possible reasons for how members of protected groups are treated. In this respect, the problem of proof for unconscious discrimination is essentially no different in kind from the problem of proving discriminatory disparate treatment of the old fashioned, deliberate kind. Because people rarely admit to taking race into account, conscious discrimination is often covert or hidden from view. Proving discrimination therefore comes down to a process of elimination. The key is to rule out other explanations. Although this methodological goal sounds simple and certainly has been a staple of discrimination law and practice for some time, it is too often ignored or slighted in discussions of unconscious bias in the service of depicting this form of discrimination as new, unprecedented, and revolutionary in its implications. But when it comes down to proving the actual existence of unlawful treatment, claims of unconscious discrimination present remarkably few novelties.

To see this, it is necessary to consider how one would go about demonstrating unconscious disparate treatment—in other words, proving that someone's decisions have been unconsciously influenced by another person's race or sex. There are two potential approaches.

The first is retrospective. Social life is replete with decisions that could potentially be based on a person's race or sex. The challenge is to identify the panoply of other possible non-race or sex-based decision-making variables, and to use statistical techniques and regression analysis to factor out their influence. The goal is to isolate the effects of characteristics like race or sex and to assess whether these traits actually made a difference to outcomes. The difficulties of this task are well-known. It is no easy matter to identify and measure the range of pertinent characteristics that might affect how individuals are treated. It is hard to demonstrate which factors decision-makers actually relied on and whether those actually made a difference to outcomes.

The second approach is prospective: one can design experiments to equalize non-protected factors that might affect how a particular person is treated, while varying only that individual's race or sex. Because other inputs into the decision are controlled, such experiments can show that race or sex "made the difference" to an outcome. As discussed more below, such experiments are often cumbersome, expensive, and hard to perform. Moreover, matching inputs is difficult even in a highly controlled setting, and the laboratory conditions in which these experiments are conducted do

not necessarily have significant application to the real world.

Part II of this Article will discuss some real-world situations that give rise to allegations of unconscious group-based bias. It will address the challenge of substantiating those claims based on the analysis of retrospective data. In Part III, this Article will examine social and laboratory experiments designed to determine whether social decision-making is infected by unconscious group-based generalizations. It will critically analyze attempts to link such evidence to legally actionable discriminatory treatment.

II. RETROSPECTIVE CLAIMS OF DISCRIMINATION

A number of large corporations—such as Walmart, Home Depot, Shoney's, Texaco, Coca-Cola, and FedEx—have been the targets of large-scale anti-discrimination suits in recent years.[14] In a related vein, the status of women in science and engineering careers has recently elicited growing concern. These trends stem from a common pattern: All too often, minorities and women do less well than white males in competitive spheres like business, government, or academia, as measured by pay, positions of authority, and advancement.[15] These disparities give rise to accusations of discrimination. Yet the accused companies or institutions typically deny that they harbor any bias, and strenuously profess their deepest commitment to equal opportunity, fair procedure, and neutral criteria. Many expend considerable time and effort rooting out any traces of discrimination, even to the point of engaging the services of professional "diversity consultants" to scrutinize and revise their personnel practices. The juxtaposition of denials of discriminatory motive with differential group results leads directly to claims of unconscious bias. The contention is that, although these social actors may sincerely believe they are scrupulously evenhanded and committed wholeheartedly to "equal opportunity," they are wrong! They may not realize it, but their operations are riddled with racism or sexism.

Unfortunately, such loose talk is often marked by sweeping claims that are oblivious to the methodological difficulties and intellectual challenges of substantiating them. Too many discussions of unconscious bias display a wholsesale disregard for alternative possible explanations for observed group disparities. In addition, attributions of unconscious bias too

---

[14] *See, e.g.*, Nancy Levit, *Mega-Cases, Diversity, and the Elusive Goal of Workplace Reform*, 49 B.C. L. REV. (forthcoming 2008) (manuscript at 1–3), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1017539 (discussing recent employment discrimination cases involving large corporations); Cheryl L. Wade, *The Impact of U.S. Corporate Policy on Women and People of Color*, 7 J. GENDER, RACE & JUST. 213, 223–24 (2003).

[15] *See, e.g.*, AMY CAIAZZA ET AL., INSTITUTE FOR WOMEN'S POLICY RESEARCH, WOMEN'S ECONOMIC STATUS IN THE STATES: WIDE DISPARITIES BY RACE, ETHNICITY, AND REGION 1, 6, *available at* http://www.iwpr.org/pdf/R260.pdf.

frequently conflate theories of disparate impact and disparate treatment. Complaints about seemingly neutral protocols that affect groups differently are coupled with accusations that otherwise similar or similarly qualified individuals are treated worse because of their identity. This conflation is often accompanied by a failure to think clearly about arguments appropriate to these distinct complaints.

A. *Race and Juvenile Justice*

To illustrate how distinct contentions come together and how problems of proof are approached in the context of unconscious bias claims, consider a piece by Olatunde Johnson that addresses the over-representation of minority youth in the juvenile justice system.[16] The author's critique sounds in both disparate impact and disparate treatment, with contentions about unconscious and inadvertent bias interspersed throughout. At times, the author seems to assert that similarly situated black youthful offenders are treated more harshly than whites. Here she looks at whites and blacks that match on particular characteristics: black youth are detained more often than whites even if they possess a comparable history of delinquency and have committed the same offense.[17] The implication is that this differential treatment is unjustified. Since there is no good reason for these observed patterns other than race, she implies, then race must be making the difference.[18] It does not follow, however, that juvenile justice officials know they are discriminating. They may sincerely believe they are fair and even-handed and may be unaware that they are treating black youth more severely. Nonetheless, officials are likely influenced by stereotypes and generalizations about black youthful offenders. Although these allegations amount to claims of unconscious bias, they sound in disparate treatment.

Elsewhere the author broadens her focus to acknowledge that observed racial disparities might be due, at least in part, to the use of a broad range of neutral criterial that go beyond the narrow characteristics of the offender and offense at issue.[19] Officials may base their disposition decisions on background factors that are thought to correlate with recidivism and danger to the community. But application of these established procedures results in minority youth receiving harsher treatment than whites. They are more often incarcerated or detained and less often released to their family's custody and supervision. The criteria at issue are race neutral, but operate to the disadvantage of blacks. Claims about unconscious bias come into

---

[16] Olatunde C.A. Johnson, *Disparity Rules*, 107 COLUM. L. REV. 374, 379 (2007).

[17] *Id.* at 403–04.

[18] *See id.* at 383 (suggesting that black offenders "are given harsher, more restrictive dispositions than are their white counterparts, when controlling for relevant factors such as the offense committed").

[19] *Id.* at 404–05.

play here, too. Johnson's contention seems to be that such seemingly neutral structures are created and perpetuated at least in part because people welcome, or at least are indifferent to, the differential impact on minority populations.[20] In this, she takes some inspiration from Charles Lawrence and others in advancing something akin to a negligence theory. Although these contentions are open to interpretation,[21] they mainly sound in disparate impact. Such disparate impact claims—which focus on the differential group effects of neutral criteria—are distinct and conceptually different from claims of disparate treatment, which are directed at decisions taken *because of* race.

In neither case, however, is the unconscious element the crux of the matter. Rather, the key issues are familiar. Is the system defective in the ways Professor Johnson contends? What should be done about it? Consider the charge of disparate impact. In a laudable burst of candor, Johnson concedes that the practice of weighing such factors as family background, neighborhood of origin, and available social supports in deciding whether to detain or release juvenile offenders operates to the disadvantage of minority youth. Black youthful offenders are more likely to come from single parent homes. Their parents may, on average, be less able to offer adequate support and supervision to wayward children. In addition, given their socio-demographic situation, black offenders' parents may, on average, be more difficult to contact and less likely to show up for interviews.[22] They may be perceived—or actually be—less willing to work cooperatively with the system. If these differences exist, black youth will indeed end up being detained more often.

Professor Johnson strongly suggests that this differential impact is undesirable and should be corrected. How should this suggestion be evaluated? Her recommendation raises important normative and empirical questions. On the normative side, the first step is to identify the goals of the system and assess whether those goals are legitimate. On the empirical side, the question is whether the existing framework advances legitimate goals and whether changing it to minimize disparate impact will undermine important objectives.

Clearly, the central aim of the existing system is to promote public safety by reducing the risk of youthful recidivism. Does the current system, with its stress on family cooperation and supervision, actually work to minimize recidivism? Will changing standards for detention to reduce racial differentials result in juvenile offenders committing more crimes? Will this expose vulnerable communities to greater danger and ultimately undermine their interests?

---

[20] *See id*. at 382–85.
[21] See *supra* pp. 2–3 for a discussion on this ambiguity.
[22] Johnson, *supra* note 16, at 405.

Critical to answering these questions is the process of validation. The challenge is to show that a neutral criterion is related to or predictive of a desired outcome. To this extent, there is nothing new here. The battle of validation for criteria with disparate impact has been a staple of anti-discrimination law since the beginning. And that demand has little to do with conscious, as opposed to unconscious, motives or processes. Perhaps it can be argued that there is indeed something new: obliviousness to racial impact may enhance the risk that neutral rules harmful to minorities will be adopted without regard to whether they advance any valuable purpose. But even this argument ultimately rests on whether the criteria employed are in fact socially justified.

In the case of juvenile justice, the key empirical questions are whether, and to what extent, differences in family background and cooperativeness actually predict recidivism. Will less restrictive dispositions extract a price in greater danger from youthful offending? But even if reducing the detention rate of minority youth would come at the expense of public safety, the question remains of whether the tradeoff is worth it. What is the magnitude of the gain relative to the loss, and how should costs and benefits be compared? Does it matter that the victims of youthful recidivism are likely to be members of minority groups as well? In the end, we cannot avoid wrestling with these tradeoffs. But adding the element of unconscious bias does not change the calculus. What matters is the bottom line need to balance differential racial impact and public safety. In that vein, Professor Johnson herself acknowledges that family cooperation and family supports may be relevant to the success of so-called "diversion programs" designed to keep juvenile offenders out of detention.[23] She nonetheless recommends that less weight be put on such background factors regardless of their potential predictive value. There are, however, no easy answers here. Depending on the actual facts, others might strike the balance differently.

But what of Professor Johnson's claims of disparate treatment—the assertion that juvenile justice officials are treating offenders differently *because of* their race? Justifying this contention requires demonstrating that officials are in fact relying on the race of the offender in determining dispositions. And the necessary showing is the same regardless of whether that reliance is deliberate or inadvertent. How would one prove such reliance? Ideally, one would demonstrate that black offenders are treated worse than white offenders *who are equivalent in all other respects pertinent to disposition.* That is, blacks are given harsher sentences even when matched on all the attributes and background characteristics that officials take into account in deciding what to do with juvenile offenders.[24]

---

[23] *Id.* at 405.

[24] In contrast with disparate impact, claims of disparate treatment do not depend on whether or not

Obviously, this showing is not made merely by noting that blacks get harsher treatment. They may *deserve* harsher treatment, at least on the criteria that the system deems relevant. Likewise, it cannot be assumed that the populations of white offenders and black offenders are equivalent overall. Specifically, blacks and whites could well differ on the sociological, behavioral, and demographic characteristics that determine disposition within the system.

In fact, there is no reason to expect that the populations of black and white juvenile delinquents are similar in all pertinent respects—and good reasons to believe they are not. Various offenders may differ by race in family structure, family background, and neighborhood characteristics. Given what we know about the racial incidence of single parenthood, multipartnered fertility, educational failure, and unemployment,[25] it is entirely possible that black delinquents on average come from less orderly and intact homes or lack a resident, known, or employed father. These characteristics are established risk factors for anti-social behavior.[26] Likewise, blacks might disproportionately come from crime-ridden or gang-infested neighborhoods, be at greater risk for involvement in drug-dealing, or have a close relative or father with a criminal record. Because these factors predict greater involvement in unlawful activity, black juveniles within the system may on average commit more severe infractions or have more extensive prior records than whites. But group differentials in predictive background factors could hold even for black and white offenders with equivalent records who commit similar crimes. Systematic socio-demographic disparities can make race an independent predictor of future behavior.

That people do not come matched on all characteristics other than the forbidden one of race makes proof of discrimination more difficult. Showing that individuals have in fact been judged by race requires isolating the influence of race by controlling for all other relevant factors. This is easier said than done. Although the process of deciding what to do with juvenile offenders seems cut and dry, it is not. That offenders come in many types and varieties, and from many different backgrounds and situations, makes individual comparisons problematic. It is hard to take the full measure of each person who has broken society's rules. Thus, the concept that like cases should be treated alike sounds good in theory, but is difficult to apply in practice.

---

decision-makers are justified in taking particular non-racial criteria into account. That is, the question of whether decision-making criteria are predictive of anything the system should care about is relevant to justifying disparate impact but not to proving discrimination *because of* race. Even if the other criteria actually relied upon are arbitrary and irrelevant to recidivism, actionable disparate treatment occurs only if race influences outcomes.

[25] For discussion of family structure and race, see *infra* pp. 19–21.

[26] *See infra* notes 44–52 and accompanying text.

The fact of systematic group differences makes demonstrating unlawful discrimination even more problematic, because reliance on valid criteria can mimic reliance on race. It is unclear what criteria juvenile justice personnel actually take into account. Maybe officials give weight to characteristics like background, personal attributes, or individual behavior—but maybe they do not. Even if harsher treatment of black youth seems actuarially justified based on their actual behavior, that does not rule out the possibility that officials are discriminating based on race. The reason is that race may itself be a valid proxy for, or predictor of, future dangerousness. If race correlates with recidivism, for example, we can never be sure whether officials are engaging in "statistical" or "rational" discrimination based on race—that is, reliance on race where race is a valid predictor—or are in fact responding directly to neutral criteria that merely correlate with race.[27] Using race in the decision-making process—whether "rationally" or not—is unlawful disparate treatment, while using criteria that merely correlate with race is not. Nonetheless, these are difficult to tease apart, especially where information relied on may not be readily observable to social scientists or outsiders. Statistical discrimination—that is, reliance on race as a valid predictor—is notoriously hard to detect and eradicate precisely because it can create patterns that would be expected from reliance on valid, neutral characteristics.

Nonetheless, the observation that disparate outcomes do not necessarily mean discrimination—in the sense of disparate treatment based on forbidden characteristics—should not be forgotten. That observation highlights a reality that many do not want to face. Populations often differ in ways that might legitimately bear on how they are treated by others. Specifically, there exist systematic behavioral differences by race that have predictable social consequences. Once again, to show unlawful disparate treatment, it is necessary to show that people who match in all relevant respects are nonetheless treated worse. Where populations systematically differ in characteristics that bear on social outcomes, this a tall order.

B. *Sex and Science Careers*

Similar points are in order regarding accusations of unconscious gender bias in the realm of employment and careers. Consider, for example, this passage from a recent National Academy of Sciences report on the status of women in science: "It is not lack of talent [that causes the disparity with men], but unintentional biases and outmoded institutional structures that are hindering the access and advancement of women."[28]

---

[27] For references and discussion of "rational" discrimination, see *infra* p. 32 & n.83.

[28] COMM. ON MAXIMIZING THE POTENTIAL OF WOMEN IN ACADEMIC SCI. & ENG'G, NAT'L

This statement bears closer examination. In identifying "unintentional" bias as a moving force behind gender disparities in scientific careers, the main focus of this statement is on institutional structures. Conventions, procedures, and criteria for assigning rewards in science are suspect because they operate to impede women's advancement. This is a garden-variety disparate impact claim. As such, it raises the familiar question pertinent to such claims: do the criteria, conventions, and organizational dynamics at work—the settled ways of doing things—serve any useful purpose or advance valued goals? Could they as well be different without any significant cost to scientific quality, output, efficiency, innovation, or anything else we value or care about? The National Academy of Sciences Report seems to suggest that these settled practices lack validity—that they do not serve any important institutional or social purposes. Implicit in the critique of the current system is an attack on that system as arbitrary or unnecessary.[29]

This passage could also be read as asserting unconscious disparate treatment. The central contention here is that gender itself influences the calculus. Despite measuring up in every way, women are treated differently, and less well, *because of* their sex. Although the claims implicit in the NAS statement (disparate impact and disparate treatment) are not strictly inconsistent—as both could characterize the system simultaneously—they are somewhat in tension. The first concedes that women may differ from men—although in ways that have no bearing on quality or productivity—while the second claim is that women are treated worse despite being the same as men.

The critical part of the NAS statement for our purposes is the contention that women's relative lack of success is due to "unintentional biases." Here, the concerns with unconscious bias and evaluative structures come together. *Ad hoc*, subjectively loaded, unsystematic methods for assessing performance permit an illegitimate consideration—specifically, the sex of the target—to creep into the evaluative process. How should this contention be approached? Assessing it requires the same analysis as claims of illegitimate bias on the job more generally. Ability, merit, productivity, work product, and performance are notoriously

---

ACAD. OF SCIS., NAT'L ACAD. OF ENG'G, AND INST. OF MED., THE NAT'L ACADS., BEYOND BIAS AND BARRIERS: FULFILLING THE POTENTIAL OF WOMEN IN ACADEMIC SCIENCE AND ENGINEERING 1 (2007) [hereinafter BEYOND BIAS AND BARRIERS].

[29] For a discussion of the challenges posed by the push to make science more hospitable to women, see Christina Hoff Sommers, *Why Can't a Woman Be More Like a Man*, THE AMERICAN, March–April 2008, at 31, *available at* http://www.american.com/archive/2008/march-april-magazine-contents/why-can2019t-a-woman-be-more-like-a-man (criticizing the idea that women's success in science "depends on changing the rules of the game"). *See also* SUSAN PINKER, THE SEXUAL PARADOX: MEN, WOMEN AND THE REAL GENDER GAP (2008) (discussing how psychological gender differences might lead more men than women to choose and achieve in science).

resistant to precise measurement. Fair evaluation requires difficult comparisons and the exercise of discretionary judgment. As applied to the real-world workplace, the assertion that such assessments are tainted by race or sex is easy to make but hard to prove or refute.

Those who would challenge the use of subjective elements in workplace evaluations on the grounds of their vulnerability to distortion by unconscious stereotypes must offer an objective baseline against which such distortion can be measured. Without such a benchmark it is impossible to show that like cases are not treated alike. Reliable measures of productivity or quality may be available if tasks are straightforward and amenable to direct observation or precise quantification. But many modern jobs are complex and require the exercise of discretion and judgment. There are no cut and dried criteria for productivity and no straightforward yardsticks for success. Because this hobbles attempts to prove that bias is distorting the process as well as efforts to show that it does not, it matters who is required to bear the burden of proof. But this means that requiring complex organizations to prove that all personnel practices are objectively fair is a heavy burden indeed and even a formula for paralysis. The demand also begs the question of what procedures will replace existing ones, and whether they can be expected to improve the situation.[30]

Is there some way around this dilemma? Purely objective measures of on-the-job performance are just not feasible for many jobs, or come at too great a cost to the effectiveness of the enterprise. The initial hiring decision is potentially more tractable, if only because less information is available about workers at this stage and objective measures work as well as any. Such screening devices were once commonly used in deciding whom to hire. General ability testing, for example, is an effective—albeit imperfect—method for selecting workers in a broad range of jobs, because general cognitive ability—of the type measured on IQ tests—is the best known predictor of employee performance regardless of job complexity.[31] Yet objective instruments that emphasize cognitive aptitude have long been attacked for their disparate impact on some minorities and, in some cases, women.[32] In the wake of these objections, such screening devices

---

[30] *See, e.g.*, Samuel R. Bagenstos, *The Structural Turn and the Limits of Antidiscrimination Law*, 94 CAL. L. REV. 1, 17–26 (2006) (noting the pitfalls of proposals for reforming workplace practices); Wax, *supra* note 2, at 1139–42 (questioning whether efforts to identify and eliminate unconscious bias will actually prevent disparate treatment).

[31] *See* Mark Kelman, *Concepts of Discrimination in "General Ability" Job Testing*, 104 HARV. L. REV. 1157 (1991) (discussing general aptitude job testing as a suspect method of predicting job performance); 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 the "overwhelming research evidence showing a strong link between general cognitive ability . . . and job performance").

[32] *See* Kelman, *supra* note 31, at 1159, 1240 (detailing the group disparate impact of "general ability" job testing).

have been virtually eliminated from human resources practice and replaced by factors such as educational credentials, job experience, personal interviews, and previous performance. Yet because some of these criteria are relatively *ad hoc* and subjective, they are also open to attack as potentially vulnerable to distortion by group-based stereotypes and biases.

C. *Race and Employment*

As with sex, so too with race. Although recent decades have seen steady progress, there are still pronounced disparities by race, especially for men, in job status, earnings, and rates of employment.[33] These patterns have provided the occasion for decades of job discrimination lawsuits against public and private employers. Once again, the emphasis is on disparate outcomes. The notion that racial differentials are sufficient to create a presumption of discrimination in employment is a staple of legal doctrine and practice in this field. The traditional doctrine also provides, however, that the presumption is not conclusive. It can be rebutted by a showing that observed differences reflect disparities in qualifications or performance.

Longstanding understandings surrounding the significance of workplace disparities have come under pressure from the recent stress on unconscious motivation. Common job evaluation practices, it is claimed, are inherently vulnerable to unconscious bias. Organizations must confront the fact that their practices are subject to distortion despite avowed commitments to equal opportunity and strenuous efforts to be evenhanded and fair.[34] Unfortunately, this rhetoric has tempted some legal scholars interested in unconscious bias to treat mere possibility as established fact. That common organizational practices are, by definition, tainted has acquired the status of an unquestioned proposition. And accepting that proposition is seen to justify the inference that unconscious discrimination is the source—and the only source—of observed racial differences in the success of individuals within such organizations.[35] In the context of scholarship on unconscious bias, the traditional presumption that attached to racial disparities has thus morphed into a firm—and irrebuttable—conclusion: Unconscious discrimination is the sole source of

---

[33] *See, e.g.*, Finis Welch, *The Employment of Black Men*, 8 J. LAB. ECON. S26, S27–S30 (1990) (analyzing census employment data).

[34] *See generally* Bagenstos, *supra* note 30 (describing such arguments); Susan Sturm, *Second Generation Employment Discrimination: A Structural Approach*, 101 COLUM. L. REV. 458 (2001); Phillip E. Tetlock & Gregory Mitchell, *Implicit Prejudice and Accountability Systems: What Must Organizations Do to Prevent Discrimination?*, RES. ORG. BEHAV. (forthcoming 2008) (on file with Connecticut Law Review).

[35] *See* Bagenstos, *supra* note 30, at 5–8 (noting theories of how unconscious bias affects workplace inequality); Sturm, *supra* note 33, at 468–74 (labeling unconscious bias "second generation" discrimination).

differential outcomes by race.

This position is seriously open to challenge. Specifically, it ignores significant "supply side" differences by race. The expectation of equal performance fails to grapple with the realistic possibility, grounded in social fact, that workers from different groups may not perform equally well overall. And this can be so even when workers have been hired based on similar credentials of the type often relied on in the employment setting. Although conventional types of job screening methods would seem to create a pool of workers equally capable of performing the job, that is not necessarily so. As with juvenile justice and women in science, so with employment more generally: even with populations that are matched on many characteristics, group differences in workplace performance could easily emerge.

The possibility of group job performance disparities based on so-called "supply side" factors is one of the verboten subjects, the unspoken elephants in the room, in unconscious bias discourse. Employees hired into identifiable job categories from different groups may not be similarly qualified and endowed with attributes that enable them to perform equally well on the job. Indeed, there is considerable social scientific evidence that, at least when it comes to race, they may not, at least for now, be equally endowed.

The following hypothetical case scenario illustrates this point.[36] Floret, Inc. is a large nationwide distributor of flowers. The company purchases and arranges flowers, plants, and bouquets, and then delivers them to florists and consumers nationwide. As part of its distribution team, Floret employs hundreds of truck drivers. The driver's job is to deliver customer orders to the right place within the promised time frame. A premium is placed on accuracy, speed, safety, and efficiency. In pursuing these goals, drivers have discretion in planning routes and dealing with contingencies that arise during the delivery process.

How does Floret go about hiring drivers? The job qualifications are a high school diploma or G.E.D. diploma, or, in lieu of that, honorable discharge from military service. A valid truck driver's license and clean driving record are required. A few years of job experience are also highly desirable. A company personnel officer conducts a brief interview and peruses the names of—but does not always personally contact—references provided by the job candidate. Candidates must disclose any arrests, convictions, and prison terms, although not all brushes with the criminal justice system are considered *per se* disqualifying.

---

[36] This scenario is modeled on an ongoing case with which the author is familiar. Because the identities of the parties and the details of the litigation are confidential, the author has disguised or altered some aspects of the case. With respect to the basic elements of the argument, however, the account accurately reflects the facts.

Once drivers are hired, Floret regularly evaluates their on-the-job performance. Because some, although not all, aspects of performance are fairly transparent and can be easily assessed, some evaluative criteria are objective. Drivers are rated on absenteeism, tardiness, number of orders delivered, errors in delivery, timeliness of delivery, and driving safety. But drivers are also rated on "soft" aspects of performance. Scores are given for cooperativeness, responsiveness to direction, flexibility, teamwork, work attitudes, ability to work independently, leadership, innovative thinking, and the like. These measures are assessed by direct supervisors on the basis of observation and day to day experience. However, all dimensions are rated quantitatively in keeping with the best human resources practice.[37]

Unfortunately, assessments of Floret drivers, according to these procedures, did not produce uniform results. Black drivers, who, like all the drivers, were overwhelmingly male, received lower average scores than whites and were more likely to be disciplined, laid off, or fired. White drivers were promoted to supervisory positions at significantly higher rates than black drivers and received more pay raises. In response to these differentials, the black drivers joined together to file an employment discrimination lawsuit against Floret. Even though black drivers, on average, got worse marks on the objective as well as the "soft" parts of their evaluations, the plaintiffs decided not to challenge all aspects of the assessment process. Rather, they confined their challenge to the subjective measures as described above.[38] Their contention was that the very use of these "soft" criteria violated anti-discrimination laws because (1) these measures have a disparate impact on black employees, without any showing that the ratings categories are job related or predictive of actual performance—an argument that sounds in disparate impact—and (2) these

---

[37] *See, e.g.*, Robyn Dawes et al., *Statistical Prediction vs. Clinical Prediction: Improving What Works*, *in* A HANDBOOK FOR DATA ANALYSIS IN THE BEHAVIORAL SCIENCES: METHODOLOGICAL ISSUES 351, 352–60 (Gideon Keren & Charles Lewis eds., 1993) (gathering studies showing that assigning a quantitative value even to subjective or soft criteria maximizes the predictive validity and internal consistency of assessments).

[38] Indeed, the decision not to challenge the objective criteria is not surprising in light of the fact that the drivers' objective scores correlated closely with the subjective ratings, a pattern that suggested that the subjective assessments were not significantly biased. Because the Floret plaintiffs declined to challenge the objective measures, these facts would not ordinarily be brought out in the case.

The Floret data mirrors the high correlation generally reported in the human resources literature between objectively verifiable measures of performance and more open-ended subjective criteria. This correlation undermines the assumption, central to the unconscious bias literature, that such subjective job evaluations are fatally distorted by category-based biases. *See* H. W. Hennessey, Jr. & H. John Bernardin, *The Relationship Between Performance Appraisal Criterion Specificity and Statistical Evidence of Discrimination*, 42 HUM. RESOURCE MGMT. 143, 155 (2003) (concluding that there is only "the very slightest support . . . that greater criterion specificity . . . is related to less adverse impact against minorities, females, or older workers"); Philip L. Roth et al., *Ethnic Group Differences in Measures of Job Performance: A New Meta-Analysis*, 88 J. APPLIED PSYCHOL. 694, 701–02 (2003) (analyzing objective versus subjective measures of performance).

criteria are prone to distortion—and indeed, are destined to be distorted—by unconscious, race-based biases and stereotyping—an allegation of disparate treatment. The plaintiffs argued, in effect, that the very subjectivity of the categories for rating drivers allowed unconscious biases to creep in and virtually guaranteed that blacks would be evaluated less fairly and more harshly than whites despite similar performance.

In response to these allegations, Floret denied all charges. The company strenuously expressed its commitment to racial fairness, equal opportunity, and respect for diversity. It pointed to investments of time and effort in developing policies designed to reach out to and recruit minorities, achieve a diverse and culturally sensitive workforce, and evaluate workers in an even-handed manner.

The Floret case has many features common to claims of unconscious bias. The employees were assessed on "soft" criteria, which are deemed vulnerable to distortion by unconscious stereotyping. Workers appeared similarly qualified at the outset, thus begging the question of why the system produced racially disparate results. And the defendant denied discriminating against minority workers, pointing to elaborate safeguards designed to ensure an evenhanded process. In keeping with the dominant understandings in the legal scholarship on unconscious stereotyping, the plaintiffs argued that none of this really matters. Unconscious bias can operate unbeknownst to the operators, and despite their best intentions.[39] A company need not be aware that their processes are structurally infirm and discriminatory. Those defects will occur despite the company's best efforts.

Once again, the critical but tacit assumption underlying the Floret lawsuit is that observed racial disparities are due to discrimination. Once again, the problem with this inference is that it is unjustified. It flies in the face of a substantial body of social science data documenting racial differences in "supply side" dimensions that bear on performance. As already noted, this literature supports a prediction that blacks and whites will not always perform equally well on the job, even when workers from these groups are matched on credentials that employers frequently use to screen and hire job candidates. Specifically, the data provides reason to believe that black and white Floret drivers will not perform equally well. Although the background information does not definitively resolve the question of whether there are differences in these particular drivers' performance by race, it sets up a rival explanation to discrimination as the possible source of observed results.

What social science evidence is pertinent here? Since Floret demands a high school diploma, one might ask whether all high school graduates are

[39] For insight into this point-counterpoint, see Tetlock, *supra* note 23.

equally qualified. Specifically, is there reason to believe that black and white high school graduates, on average, differ on dimensions that employers might care about or that actually correlate with and predict performance on a job like driving a delivery truck for Floret? The answer is yes. First, the average black high school graduate leaves high school with significantly lower academic—math and reading—skills than the average white graduate. As measured by scores on the National Assessment for Educational Progress (NAEP), the average 17-year-old black student possesses the math and readings skills of the average 13-year-old white student—a stark difference.[40] Are these types of academic skills relevant to driving a delivery truck? Almost certainly they are. As already noted, a large body of empirical work indicates that general cognitive aptitude predicts job success.[41] Although partly a function of native ability, such aptitude is developed through education and is strongly reflected in academic achievement. Plain common sense also suggests that basic reading and math skills are capacities that drivers need and will draw on routinely. Drivers must read maps and delivery orders, plan routes, cope with traffic problems, interact with superiors and customers, accommodate last minute changes and demands, and deal with mix-ups and incomplete information. These tasks require the daily use of skills—both cognitive and non-cognitive—that contribute to, and are reflected in, academic achievement.[42]

In addition to measured differentials in skills and learning, there are other characteristics of the populations from which drivers are drawn that would predict average group disparities in job performance. There is now a large literature showing that family structure of upbringing affects

---

[40] NAT'L CTR. FOR EDUC. STAT., U.S. DEP'T OF EDUC., THE NATION'S REPORT CARD, NAEP 2004: TRENDS IN ACADEMIC PROGRESS: THREE DECADES OF STUDENT PERFORMANCE IN READING AND MATHEMATICS 33, 42 (2004), *available at* http://nces.ed.gov/nationsreportcard/pdf/main2005/2005464.pdf.

[41] *See generally* Schmidt, *supra* note 31 (documenting research evidencing a "strong link" between general cognitive ability and job performance).

[42] On the relationship between non-cognitive skills, academic achievement, and job success, see, e.g., David Autor & David Scarborough, *Will Job Testing Harm Minority Workers?* 6–8 (M.I.T. Dept. of Econ., Working Paper No. 04-29), *available at* http://ssrn.com/abstract=580941 (presenting data on personality tests administered by one large employer that reveal ethnic group differences in important job-related attributes like agreeableness and conscientiousness); Pedro Carneiro & James J. Heckman, *Human Capital Policy*, *in* INEQUALITY IN AMERICA: WHAT ROLE FOR HUMAN CAPITAL POLICIES? 77, 92–93, 100–01 (Benjamin M. Friedman ed., 2003) (discussing the contribution of both cognitive and non-cognitive skills to success in the labor and education market); Angela L. Duckworth et al., *Grit: Perseverance and Passion for Long-Term Goals*, 92 J. PERSONALITY & SOC. PSYCHOL. 1087, 1098–99 (2007), *available at* http://www.sas.upenn.edu/~duckwort/images/Grit%20JPSP.pdf (reporting that "grit," perseverance, and focus on task, as well as personality traits like conscientiousness and agreeableness, can be shown to contribute significantly, over and above cognitive ability, to academic achievement and job success); Lewis R. Goldberg et al., *Demographic Variables and Personality: The Effects of Gender, Age, Education, and Ethnic/Racial Status on Self-Descriptions of Personality Attributes*, 24 PERSONALITY & INDIVIDUAL DIFFERENCES 393, 384, 401 (1998) (finding average personality differences between whites, blacks, and Hispanics in a large national sample).

behavioral outcomes, including, but not limited to, school success. In particular, the setting in which children grow up correlates with many aspects of effective socialization and future behavior, and those effects are observed even for families with comparable income and parental education. For example, social scientists have documented that children from non-traditional families have an enhanced risk of problems in many spheres of life.[43] Children living with single parents are significantly more likely to drop out of school, experience unemployment, and have an out-of-wedlock child themselves.[44] They also have "lower educational attainment, poorer mental health, and more family instability when they grow up."[45] Recent research reveals that individuals raised in blended or step-parent families also are at risk relative to those from more traditional backgrounds. In addition to having lower educational achievement and completing fewer years of schooling, these persons experience relatively more behavioral and psychological problems and have less stable adult relationships.[46] Indeed, children from blended families fare no better than children raised by single or divorced parents.[47] In sum, data from a variety of sources now strongly suggest that children growing up in settings other than traditional families are at a disadvantage.[48]

How is this relevant to the Floret case? The traditional nuclear family is much less common among blacks than other major American ethnic groups. Dramatic family structure differences by race are reflected in out-

---

[43] *See* Kristin Anderson Moore et al., *Marriage from a Child's Perspective: How Does Family Structure Affect Children, and What Can We Do About It?* Child Trends Research Brief (June 2002), http://www.childtrends.org/files/MarriageRB602.pdf.

[44] David T. Ellwood & Jonathan Crane, *Family Change Among Black Americans: What Do We Know?*, 4 J. ECON. PERSPS. 65, 70 (1990).

[45] Sara McLanahan, *Diverging Destinies: How Children are Faring Under the Second Demographic Transition*, 41 DEMOGRAPHY 607, 611 (2004).

[46] *See* ABIGAIL THERNSTROM & STEPHAN THERNSTROM, NO EXCUSES: CLOSING THE RACIAL GAP IN LEARNING 132 (2003); McLanahan, *supra* note 45, at 611; Moore et al., *supra* note 43, at 1–2; Wendy Sigle-Rushton & Sara McLanahan, *Father Absence and Child Well-Being: A Critical Review*, *in* THE FUTURE OF THE FAMILY 116, 116–125 (Daniel Patrick Moynihan et al. eds., 2004).

[47] *See, e.g.*, KATHRYN EDIN & MARIA KEFALAS, PROMISES I CAN KEEP: WHY POOR WOMEN PUT MOTHERHOOD BEFORE MARRIAGE 215 (2005) (noting that "*living apart from either biological parent at any point during childhood is what seems to hurt children*"); Donna K. Ginther & Robert A. Pollak, *Family Structure and Children's Educational Outcomes: Blended Families, Stylized Facts, and Descriptive Regressions*, 41 DEMOGRAPHY 671, 687 (2004); Sandra L. Hofferth, *Residential Father Family Type and Child Well-Being: Investment Versus Selection*, 43 DEMOGRAPHY 53, 74–75 (2006).

[48] Indeed, a picture has gradually emerged of the traditional nuclear family—consisting of two married parents living with their shared biological children—as the "gold standard," or the most desirable setting for raising children. A research brief by *Child Trends* sums up the scholarly consensus:

> [R]esearch clearly demonstrates that family structure matters for children, and the family structure that helps them most is a family headed by two biological parents in a low-conflict marriage. Children in single-parent families, children born to unmarried mothers, and children in stepfamilies or cohabiting relationships face higher risks of poor outcomes.

Moore et al., *supra* note 43, at 6. *See also* McLanahan, *supra* note 45, at 611; Amy L. Wax, *Traditionalism, Pluralism, and Same-Sex Marriage*, 59 RUTGERS L. REV. 377, 402–06 (2007).

of-wedlock birth rates, with 69% of black children now born to single mothers as compared to about 33% of white children.[49] Indeed, there are few contrasts in social life as stark as the disparity in single parent families by race, which exists across all income and education levels.[50] Likewise, divorce rates are higher among blacks than whites—a gap that also cuts across all social classes.[51] Additionally, multipartnered fertility—the practice of men and women having children by more than one partner—is significantly more common among blacks than other American ethnic and racial groups, and is more likely among blacks to result in an extra-marital birth.[52]

These trends, and the racial fault lines they reveal, indicate that blacks are significantly more likely to be brought up in single-parent, fatherless, non-marital, or blended families than whites. These patterns have prevailed for some time. To the extent these circumstances of upbringing are associated with lower academic achievement and higher risks of anti-social behavior, impulsiveness, involvement with drugs, unstable personal relationships, and mental health problems, differences in family structure alone would predict that black men are more likely to experience those difficulties. It would not be unexpected for the effects to spill over into performance on the job.

There are yet other socio-demographic differences between black and white workers that could affect productivity on the job. Just as blacks and whites have disparate family backgrounds, their adult family situations also fail to match. Over the past fifty years, marriage rates have declined precipitously among blacks, with the percentage of adults married, or ever married, now by far the lowest among major American groups.[53] This trend is particularly pronounced among black men, who marry much less

---

[49] *See, e.g.*, Ellwood & Crane, *supra* note 44; *see also* Amy L. Wax, *Engines of Inequality: Race, Class, and Family Structure*, 41 FAM. L.Q. 567 (2007).

[50] *See, e.g.*, Michael Winerip, *In Gaps at School, Weighing Family Life*, N.Y. TIMES, Dec. 9, 2007, § 14LI, at 4, *available at* LEXIS, News Library, NYT file (noting that only 35% of black children currently live with two parents, compared with 68% of American children overall).

[51] *See* Megan M. Sweeney & Julia A. Phillips, *Understanding Racial Differences in Marital Disruption: Recent Trends and Explanations*, 66 J. MARRIAGE & FAM. 639, 648 (2004).

[52] *See, e.g.*, Karen Benjamin Guzzo & Frank F. Furstenberg, Jr., *Multipartnered Fertility Among American Men*, 44 DEMOGRAPHY 583, 591–92 (2007); Wax, *supra* note 49.

[53] *See, e.g.*, Ellwood & Crane, *supra* note 44, at 68–69 (documenting and discussing racial differentials in marriage rates); Robert D. Mare & Christopher Winship, *Socioeconomic Change and the Decline of Marriage for Blacks and Whites*, *in* THE URBAN UNDERCLASS 175, 175 (Christopher Jenks & Paul E. Peterson eds., 1991) (same); R. Kelly Raley, *Recent Trends and Differentials in Marriage and Cohabitation: The United States*, *in* THE TIES THAT BIND: PERSPECTIVES ON MARRIAGE AND COHABITATION 19, 23 (Linda J. Waite et al. eds., 2000) ("Since the 1950s, black women's marriage rates have declined much more steeply than white women's."); Joy Jones, *Marriage Is For White People*, WASH. POST, Mar. 26, 2006, at B01, *available at* LEXIS, News Library, WPOST file (stating that "[t]he marriage rate for African Americans has been dropping since the 1960s, and today [African Americans] have the lowest marriage rate of any racial group in the United States. In 2001, according to the U.S. Census, 43.3 percent of black men and 41.9 percent of black women in America had never been married, in contrast to 27.4 percent and 20.7 percent respectively for whites").

frequently than other American men.[54] These differences are observed even among men with similar levels of education and income.[55] These marital patterns have behavioral implications. Social scientists have observed that married men are more law abiding, more sober in their habits, less involved with drugs, less prone to mental illness, and generally healthier. They also tend to work harder, attain higher job status, and earn more income.[56] Because white male adults are more likely than blacks to be married, it can be predicted that they will display these desirable traits and characteristics more frequently. Once again, it would not be surprising if these patterns carried over to make whites, on average, more successful on the job. And that pattern might well prevail for a particular job, such as Floret delivery driver.

In sum, these observations suggest that, overall, black males may not be as well socialized or well prepared as their white counterparts for the demands and expectations of the workplace. This could be so even for individuals evenly matched with regard to credentials—such as levels of education and years of schooling completed—that many employers rely on. These credentials operate as relatively crude screening devices that omit important, job-relevant information. That information could well reveal heterogeneity within otherwise matched populations. Indeed, many commonly employed job requirements fail to capture group differences that might bear on actual productivity and workplace success. As with juvenile justice, so with jobs: seemingly similar populations may be dissimilar in important respects. And focusing on select parameters or characteristics, such as type of offense and past record for juvenile offenders, or high school graduation for workers, may provide an incomplete picture of all dimensions bearing on future behavior or performance. More information about group differences helps make better predictions.

What are the implications for the Floret case? The social science literature on academic achievement and family structure strongly suggests

---

[54] Jones, *supra* note 53.

[55] Ellwood & Crane, *supra* note 44, at 76.

[56] *See, e.g.*, JOHN H. LAUB & ROBERT J. SAMPSON, SHARED BEGINNINGS, DIVERGENT LIVES 41–44 (2003) (suggesting ways in which marriage deters criminal behavior); Sanders Korenman & David Neumark, *Does Marriage Really Make Men More Productive?*, 26 J. HUM. RESOURCES 282, 303–04 (1991) (finding that married male workers are higher paid and receive higher performance ratings); Robert J. Sampson et al., *Does Marriage Reduce Crime? A Counterfactual Approach to Within-Individual Causal Effects*, 44 CRIMINOLOGY 465, 469 (2006) (concluding "marriage influences criminal behavior among men"); Avner Ahituv & Robert I. Lerman, *How Do Marital Status, Wage Rates, and Work Commitment Interact?* 27 (Inst. for the Study of Labor, Discussion Paper No. 1688, 2005), *available at* http://ftp.iza.org/dp1688.pdf (finding higher earnings as a result of marriage). As Korenman and Neumark show, married men's higher average earnings are not just a matter of a "selection" effect—that is, of the greater propensity of men with desirable attributes to marry. Korenman & Neumark, *supra*, at 304. Rather, marriage actually induces men to work longer hours and earn more.

that the average black Floret delivery driver is unlikely to be as capable, or to perform as well as, the average white driver. Indeed, in light of the social science data, it would indeed be surprising if these groups performed equally well. One could go so far as to say that the best prediction from the data available is that they would not. It follows that, given what we know, "supply side" differences are at least as plausible as unconscious bias as an explanation for racial disparities in the Floret drivers' performance, and could well be more plausible.

It may be argued that it is unfair or even offensive for an organization such as Floret to offer generalizations about group characteristics to defeat claims of discrimination—especially if these assertions indulge the stereotype of blacks as less capable workers. But such generalizations are necessary to defend against plaintiffs' accusations, which rely on generalizations as well. The lawsuit is based on data showing *average* differences by race: the accusation is that Floret rates black drivers lower as a group than whites. In defending against plaintiffs' contentions, it is appropriate—and indeed necessary—to show that alternative explanations for group differences in job performance have not properly been ruled out.

The discussion so far suggests several caveats. First, the evidence on average group differences says nothing about particular individuals. The data support generalizations about populations and point to differences between distinct groups as a whole. Outstanding or underperforming workers may exist in each category. Second, although the data would support predictions about job performance for groups overall, it does not show that the subset of employees in a given lawsuit is perfectly representative of any group. Nor does it definitively prove that workers for a particular company will exemplify the relative differences expected between populations—or a screened subset of populations—from which they are drawn. There may be other, unobserved or undocumented differences or similarities between workers from different races employed by a particular organization: For example, most black Floret drivers might just happen to be married. But facts bearing on whether particular workers are demographically representative or not are subject to empirical investigation.

Finally, supply side arguments of this type say nothing about the origins of racial disparities in qualifications, skills, performance, or job productivity. It is important to understand that noting these differences entails no claim about the existence of genetic differences between the races. Specifically, it implies no position on the questions of whether observed patterns are the outgrowth of past or present social forces—such as poverty, discrimination, disadvantage, or cultural difference—or whether they partly reflect biological factors.

In the setting of the Floret lawsuit, and in the absence of more individuating information, however, drawing inferences from the

characteristics of populations is the best we can do. Once again, such inferences are especially pertinent because the lawsuit is based on similar inferences—that is, from observed group disparities in performance evaluations to the existence of group-based discrimination. Although there is no direct, individualized evidence that the black delivery drivers exemplify the sociodemographic attributes that predict poorer performance, there is likewise no direct evidence that the black drivers are no different from whites in all pertinent dimensions and in fact perform just as well. In the same vein, however, the assertion that the ratings of black drivers are tainted by impermissible considerations of race is grounded solely in circumstantial evidence and established only by inference. Inadvertent race-based discrimination has not been—and indeed cannot be—directly observed or demonstrated under these circumstances. Although existing outcomes might be due to bias, the mere possibility is a far cry from proof. It takes a theory to beat a theory, and data to show which theory is correct. At best, the evidence presented leaves rival possibilities on the table without establishing the validity of either. Certainly, in the absence of something more, discrimination cannot be regarded as a better or more plausible hypothesis than the supply side explanations offered here.

This discussion merely reiterates a fundamental principle that has been central to anti-discrimination law for some time and long predates the preoccupation with unconscious bias. Disparities in employment should never be regarded as sufficient to demonstrate that discrimination is the cause of those disparities—either in whole or in part. The inference from differential outcomes by race or sex to unlawful discrimination requires excluding alternative possible explanations.

Those who would defend a stronger—or even a conclusive—inference in today's climate would point out that there is now a substantial body of scientific and empirical work that purports to demonstrate that individuals are routinely influenced in their thinking by racial or sex-based stereotypes and biases. Because these are so pervasive, it must be assumed that they influence the outcomes of decisions that permit their operation. Likewise, scholars rely on other types of behavioral data—such as field audit experiments—which they claim reveal the widespread influence of subtle unconscious biases. But as discussed more fully below,[57] the evidence falls far short of showing that unconscious bias can be assumed to distort social decision-making across the board or in any particular situation at issue. At best, existing data demonstrates that identity-based biases are intermittent, unpredictable, and context sensitive, and can be extinguished by circumstantial or individuating information. To the extent they can be

[57] *See infra* notes 64–66 and accompanying text.

detected, they do not always operate to the detriment of disadvantaged groups, but rather sometimes operate in their favor.[58] In other cases, their direction cannot be discerned.[59] More importantly, studies that purport to reveal biased thinking have not reliably linked these mental patterns to real-world decisions that the law cares about. Indeed, there are only a handful of studies that even attempt to establish the connection between thought and action, and all are seriously flawed.[60] Furthermore, none even purports to compare the effect—if any—of such biases relative to the types of supply side factors discussed above.

It may be objected that any group deficiencies in supply side determinants of success are themselves traceable to unlawful discrimination and racism—past, present, or both. Nothing here is to the contrary. But the observation that past or present societal discrimination may be implicated in generating supply side differences—as reflected in some groups, on average, being underprepared or underachieving relative to others—must be sharply distinguished from the assertion that a particular company, organization, or employer violates anti-discrimination laws by selecting, evaluating, or rewarding its employees on the basis of group identity. And that distinction stands regardless of whether discrimination is claimed to be conscious or unconscious, inadvertent or deliberate.

Thus, that blacks are sometimes, or even regularly, the victims of discrimination is not enough to establish that a particular organization, in adopting neutral methods of evaluating and assessing employees, must necessarily be demonstrating a disregard—whether deliberate, careless, negligent, or unconscious—for the well-being and interests of minorities or other protected groups. Employers and economic actors take employees as they find them. That racism has contributed to making some employees less qualified does not change the fact that they are less qualified.[61] Nor does it obviously follow that the proper response to such injustices is indiscriminately to level accusations of discrimination without proof, or to dispense with the need to marshal evidence. The question of what to do about the legacy of racial discrimination is highly charged and controversial. But the fact that private actors respond to existing disparities does not justify treating them as if they are responsible for those disparities in the first place. Nor does it follow that the proper reaction is to force companies to treat workers equally regardless of whether some fall short, or to promote workers from one racial group over others who

[58] *See infra* notes 89–91 and accompanying text.

[59] *See infra* notes 92–95 and accompanying text.

[60] *See infra* notes 104–06 and accompanying text.

[61] *See* Larry Alexander, *What Makes Wrongful Discrimination Wrong?*, 141 U. PA. L. REV. 149, 188 (1992) (suggesting that a surgeon who had his hands "villainously cut off" need not be allowed to continue performing operations).

perform better, or to somehow compensate for observed performance differences or pretend they do not exist.

Rectifying past or present injustice by addressing "root causes" is a fundamentally different exercise from claiming that a particular social actor is discriminating. This point holds regardless of whether the claimed discrimination is deliberate or inadvertent. Likewise, the solution to an unjust social order is not to redefine all differences as discrimination. Rather, it requires addressing the sources of disparities and correcting them. Alternatively, the situation calls for doing the hard work of explaining why disparate inputs shouldn't matter—or, to put it more concretely, why a person of one race should be rewarded and promoted over others despite doing a measurably worse job. In other words, it requires explaining why we should jettison the current way of doing business in favor of another set of rules. Any explanation inevitably will raise the question of how such rules should be justified. This brings us back to the familiar conundrums of whether and how rigorously selection practices should be scientifically validated. These issues are central to the debate over the legal status of disparate impact claims, but have—or at least should have—very little to do with allegations of disparate treatment, whether unconscious or not. Yet these questions are routinely conflated in the context of discussions about unconscious bias.[62] At the very least, there is a serious disconnect between the sociodemographic data on human capital development, behavior, and performance and the legal scholarship on unconscious bias. This renders many discussions of unconscious bias in the law and policy literature seriously defective, or at best radically incomplete.

In light of these blind spots, it is instructive to ask what would count as good evidence that unconscious discrimination had been purged from the procedures of a company like Floret, short of equal numerical representation by race. Apart from achieving quota-like workforce diversity, how could an accused organization exonerate itself of the allegation of unconsciously biased decision-making? Addressing this question would go far towards injecting a needed rigor into discussions of unconscious bias. Given the current legal scholarship in the field, it is hard to know what the answer to this question would be.

Finally, none of these points mean that retrospective analyses are never worthwhile. Nor does it follow that real-world practices can never reveal discrimination. But the most careful and methodical studies can only go so

---

[62] *See, e.g.*, Marcel C. Garaud, Comment, *Legal Standards and Statistical Proof in Title VII Litigation: In Search of a Coherent Disparate Impact Model*, 139 U. PA. L. REV. 455, 456–57 (1990) (discussing the problems of statistical analysis in establishing disparate impact claims). In fairness, confusion may stem partly from some degree of overlap in the kinds of statistical evidence that can be relevant to both disparate impact and disparate treatment claims. Despite this, the logic of the two claims, and the steps ultimately needed to prove them, are distinct.

far. Indeed, even the best research cannot establish with certainty that complex, meritocratic systems are distorted by group-based biases.

Some recent efforts along these lines illustrate potential pitfalls. As already noted, there is rising concern about the under-representation of women in science and their failure to reach top echelons.[63] Several studies have sought to uncover gender bias in the allocation of positions and rewards in scientific fields. One study, by Christine Wenneras and Agnes Wold, claims to be the "first ever analysis" of the peer review process for awarding post-doctoral scientific fellowships.[64] The authors attempted to determine why the Swedish Medical Research Council (MRC)—one of the main funding agencies for biomedical research in that country—awarded a relatively small number of grants to women scientists. The authors observed that the pattern of awards reflected the lower ratings women scientists received from the MRC review panels assigned to evaluate grant applications. According to the authors, the scores supposedly reflected the referees' judgments about the candidates' productivity and the value and importance of their research.[65]

In trying to determine whether female applicants were treated less favorably than male, the authors could not assume that the men and women candidates as a whole were similarly qualified or otherwise equally matched. Although the authors did not reveal detailed data concerning gender differences, if any, between the male and female applicants, their methods were directed at controlling for variation in potentially relevant dimensions such as training, productivity, and field of interest. The challenge was to determine the criteria by which candidates were actually judged and then to compare men and women with similar credentials.

In trying to gauge qualifications, the authors chose to create a "bibliometric" measure, or "total impact" index of productivity, based on quantity and quality of candidates' publications.[66] Using a regression analysis to control for other factors that might make a difference—such as area of research, training, and the particular committee evaluating the candidate—the authors found that men and women who matched on the authors' impact index did not receive the same MRC scores and were not funded at equal rates. This suggested that women needed to be significantly more productive than men in order to win support. The authors concluded that women fellowship candidates in their sample were discriminated against; they were treated less well solely because of their

---

[63] *See* discussion *supra* Part II.B.

[64] Christine Wenneras & Agnes Wold, *Nepotism and Sexism in Peer-Review*, 387 NATURE 341, 341 (1997).

[65] *Id.*

[66] "Impact factors" included the candidate's number of publications, whether the candidate was first author, the quality of the journals where the papers were published, the number of times they were cited, and the quality of the journals in which the citations appeared. *Id.* at 341–42.

sex.[67]

Are these observations conclusive? What are the possible flaws in such a retrospective study? By using statistical techniques, the authors attempted to account for any important input—or "supply side"—differences between men and women applicants, thereby isolating the effects of gender alone. Without directly reporting exactly how the committees made their decisions or what factors they relied on, the researchers developed an index that they believed accurately reflected the elements of productivity and quality that the evaluating committees would most care about.

The most important caveat attending such a design is that the researchers may have missed something. Perhaps there were differences between the male and female candidates that were not immediately obvious and not fully captured by the researchers' "impact factor" index and regression controls. And perhaps the committees observed and relied on those differences. For example, the authors say nothing about the content of the grant application itself. Nor do they mention the quality and content of recommendations or the graduate school record. Although they tried to assess the candidates' past projects indirectly by counting citations to published work, they had no direct measure of quality. Originality, generativity, creativity, rigor, and the overall importance of work to the field may not have been fully captured by the researchers' citation index. In light of these possibilities, the observed gender disparities are merely suggestive of discrimination, but far from conclusive.

The Wenneras and Wold results, although generating a good deal of attention, came from a single, small study in one country and involved data that is now more than ten years old. Other attempts to find clear sex-based discrimination in scientific research have found less impressive results. Tom Tregenza recently examined the refereeing process that controls which papers are accepted and published in the fields of ecology and evolution. Using a complicated design with multiple controls, he found women did not have a lower publication acceptance rate for submitted manuscripts overall and concluded that there was no evidence that the process is discriminatory. The author noted that women's papers scored lower on some criteria—including first authorship and number of authors on the paper—that can be shown to influence the chances of acceptance.[68] Although Tregenza's results show little convincing evidence of gender bias, his study is nonetheless open to some of the same questions concerning controls, statistical analysis, and study design that apply to the Wenneras and Wold study.

---

[67] *Id.* at 342–43.

[68] Tom Tregenza, *Gender Bias in the Refereeing Process?*, 17 TRENDS IN ECOLOGY & EVOLUTION 349, 349–50 (2002).

Anna Ledin and colleagues recently investigated women's lower rates of funding from the European Molecular Biology Organization (EMBO), both under its Long-Term Fellowship (LTF) and Young Investigator Programme (YIP).[69] Pursuant to their stated goal of "test[ing] whether unconscious gender bias influenc[ed] the decisions made by the selection committee," the authors introduced an important change in the selection process: they removed information about the sex of the applicant before sending files to the funding committees for evaluations.[70] The authors found that men's higher success rate for the LTF fellowships persisted. Despite sex-blinding, male scientists were still awarded more fellowships than women.[71]

In trying to understand the roots of this sex-blinded differential, the authors set out "to correlate the committee's decisions with an assumed unbiased measure: the bibliometric data for each applicant."[72] Although conceding that "the publication record was only a part of the committee's considerations," they nonetheless developed an index based on publications and citation frequency not unlike that used by Wenneras and Wold.[73] Comparisons based on the publication index yielded a mixed picture. Although the average "impact factor" for awardee females overall was slightly higher than for males, gender differences disappeared when the impact factor was re-computed using only first and last author publications.[74] This suggested that publications were weighted by a formula that was not captured by a straightforward citation count.

The authors also noted that the pool of female applicants overall was not as qualified as the men. In particular, the average female grant applicant had fewer publications and a lower citation count than the average male. Thus, although the average impact index for women's output appeared comparable to men's on some measures—reflecting, perhaps, higher quality of papers—the average male scientist had a greater *number* of publications. The authors speculate that this higher productivity alone—apart from quality measures, at least as assessed by the author's criteria—may have influenced the committee's decisions. The authors also found that the pattern of female under-productivity persisted even for women who were awarded fellowships. That the funded women scientists continued to lag behind the funded men in scientific output suggested that the fellowship committee's predictions of future productivity, as reflected in relatively fewer women receiving fellowships, were not unwarranted.

---

[69] Anna Ledin et al., *A Persistent Problem: Traditional Gender Roles Hold Back Female Scientists*, 8 EMBO REP. 982, 982 (2007).

[70] *Id.*

[71] *Id.* at 982–83.

[72] *Id.* at 982.

[73] *Id.* at 982–83.

[74] *Id.*

Likewise, for the YIP fellowships—which were generally reserved for junior professors—the authors found that, although women's publications had about the same "impact" scores as the men's, women applicants published fewer papers overall. In addition, there were other differences. The women applicants did more teaching, had lower rank, and presided over smaller labs than the male applicants.[75]

In general, the Ledin data suggests that the situation for women scientists is more ambiguous than indicated by the Wenneras and Wold results. In particular, the Ledin data reveals important "supply side" differences that could account for most or even all the gender shortfall in scientific grants. Women scientists are on average less productive than men—a finding that has been made repeatedly in academic science.[76] The Ledin study speculates that this could be due in part to women's greater family responsibilities or different priorities for time use.[77] It has also been suggested that women are hobbled by discrimination in the allocation of resources that might allow women to be more productive.[78] However, resources are both a reward for productivity and a facilitator of it. The causal relationships are thus difficult to sort out.

But perhaps the most dramatic difference between the Wenneras and Wold and the Ledin research is that the latter used a sex-blinded procedure. The finding that gender-blinding did not increase the number of fellowships awarded to women strongly suggests that the EMBO fellowship process was not biased against women. More specifically, it refutes the claim that unconscious bias was the cause of fewer grants being awarded to women.

The Ledin research also suggests that retrospective studies of procedures where the gender of the candidate is revealed to the decision-maker are a second-best method for assessing claims of discrimination. Identity-blinded processes are far superior because they definitively remove the influence of race or sex. Thus the question of whether the "impact index" in the Wenneras and Wold study took full account of all relevant differences between men and women—and all differences the committees actually relied on—can only really be settled by subtracting gender from the equation. Since the initial procedure for awarding the Swedish fellowships revealed gender, what is needed is a sex-blind re-

---

[75] *Id.* at 983–84.

[76] *See, e.g.*, Reshma Jagsi et al., *The "Gender Gap" in Authorship of Academic Medical Literature—a 35-Year Perspective*, 355 NEW ENG. J. MED. 281, 283 (2006) (finding that though women authors have shown increased representation since 1970 in three general medical journals, they still represent only 23.2% of first-time authors and 12.7% of senior authors); J. Scott Long, *Measures of Sex Differences in Scientific Productivity*, 71 SOC. FORCES 159, 174–75 (1992) (finding that female scientists have a lower rate of publication than male scientists).

[77] Ledin et al., *supra* note 69, at 984.

[78] *See, e.g.*, BEYOND BIAS AND BARRIERS, *supra* note 28, at 122 (noting that the "critical variable" in the gender difference in publication productivity is "access to resources").

assessment after the fact. This is easier said than done. The materials used to rate applicants would have to be purged of any information about sex. Those files would then have to be sent out for evaluation by experts in the field. A valid re-assessment would require multiple reviews, so many able scientists would have to be conscripted. Further, preserving anonymity would be difficult because the identity of some applicants might be apparent from other aspects of the file. Assuming all of these challenges could be met, however, the information obtained would be invaluable. If blinding produced better outcomes for women than the initial unblinded process, it would stand as persuasive evidence that women were not being assessed fairly in the first place. On the other hand, if outcomes for women were the same, it would indicate that gender differentials were not due to bias—conscious or unconscious—but rather to disparities in qualifications or other factors.[79]

So why not just adopt sex (or race) blinding across the board? As noted, identity blinding is a cumbersome process that is often hard to accomplish. And setting up a blinded and truly anonymous protocol may not be feasible in some settings. Many areas of scientific research are highly interactive, so it is hard to mask the identity even of junior persons working in a field. Finally, anonymous procedures can be criticized as leaving out important information needed to make the best choice. All of these concerns hinder attempts to introduce identity-neutrality in broad areas of social life.

III. PROSPECTIVE EVIDENCE OF UNCONSCIOUS DISCRIMINATION

Are there other possibilities for demonstrating the influence of unconscious bias on the types of decisions—in education, the job market, science, and other spheres—that the law cares about? Instead of examining the type of real-world outcomes that pose the difficulties of controlling for complex inputs, researchers have sought to isolate evidence of unconscious bias by examining decision-making prospectively. Yet others are focused on developing tests of stereotyping in mental processes or habits of thinking, with a few attempting to tie those measures to discriminatory behavior.

Common prospective methods for demonstrating discrimination—whether conscious or unconscious—are audit testing and matched résumé studies. Here the focus is on experimentally controlling all the inputs that are thought to go into a decision while varying only race or sex. In the early 1990s for example, the Urban Institute sent black and white testers

[79] *See, e.g.*, Claudia Goldin & Cecilia Rouse, *Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians*, 90 AM. ECON. REV. 715 (2000).

with similar paper credentials to apply for jobs in several urban markets.[80] They also sought to match the demeanor and personal appearance of the testers. They found that employers were somewhat less willing to interview and hire blacks than whites, although the differentials were not dramatic and there were no differences observed in some employment markets.[81] Likewise, Devah Pager has used audit studies to demonstrate that employers generally favor white over black ex-convicts with similar criminal records and measured personal characteristics.[82]

Because these studies arguably isolate the effect of race by controlling for many other inputs to the job selection process, they provide good evidence that employers sometimes take race into account.[83] Nonetheless, two points are in order. First, these studies do not definitively demonstrate the operation of unconscious or inadvertent—as opposed to deliberate and conscious—bias in employment. They thus cannot be viewed as specific methods for eliciting unconscious discrimination, as opposed to more traditional forms. Second, the patterns observed in these studies are most consistent with statistical or "rational" discrimination.[84] That testers have similar resumes and are matched on the main elements of personal deportment does not prevent employers from relying on race as a proxy for real productivity-related differences that are difficult to observe directly. Because, as already noted, blacks and whites differ in job-related skills *despite* similarity in many standard credentials, and because individualized information about new hires is limited, employers may find it cost effective to use race as a way to screen workers. For example, employers may believe that black ex-convicts differ on average from white ex-convicts in ways that bear on job performance. And they may believe that these differences exist despite superficial similarities in ex-convicts' criminal records, work histories, and years of education.

Are those beliefs justified? The observations pertinent to the juvenile justice situation and the Floret case likewise apply here. Average differences by race in academic achievement, family structure, and other

---

[80] *See, e.g.*, Michael Fix et al., *An Overview of Auditing for Discrimination*, *in* CLEAR AND CONVINCING EVIDENCE: MEASUREMENT OF DISCRIMINATION IN AMERICA 1, 18–20 (Michael Fix & Raymond J. Struyk eds., 1993). For a critique of this work, see James J. Heckman, *Detecting Discrimination*, 12 J. ECON. PERSPS. 101, 104–05 (1998) (discussing audit studies).

[81] Fix et al., *supra* note 80, at 20, 22, 24.

[82] DEVAH PAGER, MARKED: RACE, CRIME, AND FINDING WORK IN AN ERA OF MASS INCARCERATION 101, 106 (2007).

[83] *But cf.* Heckman, *supra* note 80, at 108 (suggesting that the audit studies do not perfectly match job candidates, and thus leave room for employers relying on factors other than race).

[84] Kenneth J. Arrow, *Models of Job Discrimination*, *in* RACIAL DISCRIMINATION IN ECONOMIC LIFE 83, 97 (Anthony H. Pascal ed., 1972) (discussing statistical discrimination in job markets); TIM HARFORD, THE LOGIC OF LIFE: THE RATIONAL ECONOMICS OF AN IRRATIONAL WORLD ch. 6 (2008) (discussing statistical discrimination); Glenn C. Loury, *Discrimination in the Post-Civil Rights Era: Beyond Market Interactions*, 12 J. ECON. PERSPS. 117, 123 (1998) (same); *see also* discussion *supra* Part II.A.

background factors bearing on socialization may carry over to subpopulations of men who commit crimes. Are black and white ex-convicts really comparable in their ability and willingness to perform on the job? Just as high school graduates may differ by race in crucial attributes bearing on their desirability as workers, ex-convicts may also. Audit studies that match for personal presentation and criminal record create the appearance of similar starting points, but cannot fully control for these background differences. Nonetheless, whether such differentials actually exist for ex-convicts is an empirical question that must ultimately be assessed directly. One approach might be to examine outcomes. Do black ex-convicts as a group do as well at work as white ex-convicts when matched on the criteria used in the audit studies? This would be a good test of whether employer discrimination of this type is "rational."

Nonetheless, prospective audit studies are still far more useful than retrospective analyses of real-world data because they can more reliably reveal reliance on race. In the retrospective situation, it is impossible to know for sure whether decision-makers are actually using race. There may be other attributes influencing choice that are not apparent to researchers or outside observers. In contrast, the audit situation can reveal reliance on race directly because other information presented to the decision-maker can be directly controlled. But the fact that testing studies can and do sometimes strongly suggest the operation of race-based discrimination does not change the social fact that the background populations from which the testers are drawn are *not* necessarily the same. That is, it does not negate the existence of background "supply side" differences by race. The employers in the audit studies may simply be playing the odds that the white employee testers will prove better than the black ones, even if they appear evenly matched. The evidence suggests that, in many cases, that assumption may be valid.

Although the audit studies may best be explained as revealing a form of "rational" discrimination, it is important to remember that this type of behavior is illegal.[85] Nonetheless, there is no reason to believe that employers' crude efforts to choose the most productive employees operate through unconscious—as opposed to quite deliberate—processes, and there appears to be nothing about this type of discrimination that recent work in cognitive psychology can especially illuminate. The more important point, however, is that observed patterns do not necessarily point to residual discrimination as the product of unconscious, hard-wired, and irrational features of the human psyche; rather, such discrimination as likely represents an efficiency-driven accommodation to the present social

[85] *See* discussion *supra* note 5 and accompanying text (discussing the illegality of disparate treatment under Title VII of the Civil Rights Act of 1964); *see also supra* Part II.A (explaining rational discrimination based on race).

reality. The reality is that, for now, race remains a cheap, albeit imprecise, proxy for performance-related traits that employers and others care about. But this characterization of residual discrimination as rooted in pragmatic concerns has very different implications than the view that invidious stereotyping is an automatic and universal feature of the human mind.

In dealing with statistical bias in hiring, the key question is what society should do about it. Not only is this type of discrimination unlawful, but it is also unfair and counterproductive. If social actors assume that individuals are representative of their underperforming group, then talented persons will receive fewer rewards than they deserve. This can undermine group members' incentives to distinguish themselves, thus discouraging self-development.[86] It is thus essential that the civil rights laws be vigorously enforced against discrimination of all kinds, whether "rational" or not. Nonetheless, eradicating statistical discrimination is difficult. Where race is a valid proxy for productivity-related traits employers have strong economic incentives to pay attention to it. And bias based on real group differences can be hard to demonstrate, because it mimics permissible reliance on job-related attributes.

One method for minimizing statistical bias is the use of more rigorous forms of screening—such as cognitive ability testing—that do a better job of sorting individuals. As already noted, such methods have come under attack as having a disparate impact on minority groups.[87] Another tactic is to focus on attacking root causes: indeed, the most effective way to defeat selection based on group differences is for those differences to disappear.[88] Any discussion of how best to bring this about is far beyond the scope of this inquiry. But whether the bias at issue is unconscious or not will have no appreciable impact on how to tackle the problem of racial disparities in society and the statistical biases that feed off of them.

This brings us back to the subject of this conference and this paper: the unconscious or inadvertent use of race to make decisions about individuals. The question is how to identify unconscious bias or detect unconscious discrimination, with emphasis on the *unconscious* element. Once again, does there exist hard and fast evidence of the operation of inadvertent—as opposed to deliberate or strategic—discriminatory behavior in social decision-making? What might that evidence look like?

On reflection, it is not surprising that data pertinent to this *specific*

---

[86] *See, e.g.*, Cass R. Sunstein, *The Anticaste Principle*, 92 MICH. L. REV. 2410, 2417–19 (1994) (discussing possible incentive effects of statistical discrimination); AMY L. WAX, RACE, WRONGS, AND REMEDIES (forthcoming 2008) (on file with Connecticut Law Review).

[87] *See* Kelman, *supra* note 31, at 1208–14 (questioning the validity of general ability tests). As Kelman notes, such tests are far from perfect—they also involve generalizations and screen out some persons who can do the job. *Id.* at 1224 & n.195. But all screening methods have this defect to some degree. *See also* discussion *supra* p. 14.

[88] For more on this point, see WAX, *supra* note 86.

question are in short supply. The principal reason is that states of mind cannot be directly observed. As already discussed, it is hard enough to figure out whether a decision is influenced by race. Proof is almost always circumstantial. Demonstrating that the discriminatory treatment was accidental—as opposed to deliberate—is harder yet. If social actors deny that they are relying on race, how do we show they are deceiving us? How are we to know what they are thinking? We are not privy to others' mental states. The challenge is finding a reliable check on these types of representations.

Only a few studies come close to meeting that challenge. Michael Norton and collaborators use a clever experimental design to suggest that subjects are not aware of how race and sex affect their choices. In their studies, student subjects were asked to evaluate files for college admission candidates and applicants for jobs. The college admission files presented a complex mix on a range of credentials—including grades and test scores. The job candidate dossiers presented credentials along two main dimensions—educational background and job experience. The students were then asked serially to decide whether to admit, or reject, each college applicant, or to hire, or not to hire, each job candidate into a specified position. The researchers found that, when the files presented a mixed picture—strong on some credentials, weak on others—the students were not consistent in weighting the various factors across all candidates. In fact, the race or sex of the candidate systematically influenced which credentials the students weighted as more important.[89]

The results for race and gender were surprising. When presented with white and black college admission applicants, subjects tended to weight credentials to favor blacks. For example, if the black candidate was stronger on grades than a white candidate, they ranked the black higher and later declared grades to be more important.[90] And this tendency was observed in all students regardless of how they rated on questionnaires designed to gauge racial prejudice. Even subjects who rated as "consciously" prejudiced were not "unconsciously" biased against blacks. In fact, quite the contrary: most students had internalized a strong norm of affirmative action. For the jobs, in contrast, almost all subjects valued the

---

[89] *See* Michael I. Norton et al., *Mixed Motives and Racial Bias: The Impact of Legitimate and Illegitimate Criteria on Decision Making*, 12 PSYCHOL. PUB. POL'Y & LAW 36, 40–44 (2006), *available at* http://www.people.hbs.edu/mnorton/norton_sommers_vandello_darley.pdf (finding in one study that participants favored black candidates over white candidates and inflated the value of some credentials to justify decisions in nonracial terms); Michael I. Norton et al., *Casuistry and Social Category Bias*, 87 J. PERSONALITY & SOC. PSYCHOL. 817, 820 (2004) [hereinafter Norton, *Casuistry*], *available at* http://www.cas.usf.edu/~vandello/NortonVandelloDarley-Casuistry.pdf ("[M]ale participants who select men for management positions justify that selection by inflating the importance of whichever qualification favors a male candidate.").

[90] *See* Norton, *Casuistry*, *supra* note 89, at 824 ("When the Black candidate had a higher GPA than the White candidate, participants inflated the importance of GPA . . . .").

credentials to favor the male job candidate.[91] However, the experiment was only conducted for a job—supervisor for a construction project—that was heavily male stereotyped. As the researchers point out, the result here would not necessarily extend to positions that were sex-neutral or female-typed.

The design of the Norton studies strongly suggests some influence for inadvertent bias, as revealed in the weighting function. The subjects were instructed to be as objective as possible, but forcing them to deal with complex data for many different candidates made systematic comparisons difficult. It is a plausible inference—but of course only an inference—that the subjects were unaware that race or sex skewed their judgments.

More importantly, these studies reveal that, even if unconscious identity-based bias does sometimes come into play, it does not necessarily operate in a predictable direction. Here, blacks did better than whites in some conditions and men than women in others. In a different context, the pattern could well shift. It is thus quite dangerous to infer from isolated experiments of this kind that bias is operating in any particular situation, or in one direction rather than another. Each instance of real-world behavior must be evaluated separately. Obviously this is burdensome and methodologically difficult. But the alternative is intolerable: to credit nefarious accusations of discrimination against minorities or women that may well have no basis in fact.

Is there other research that provides persuasive evidence for the operation of inadvertent biases in real-world tasks? Recent unpublished papers by two teams of economists describe intriguing attempts to detect race-based discrimination in the sports world. Unlike the Norton studies, this work is not based on prospective laboratory experiments. Rather, it makes use of retrospective data.

Joseph Price and Justin Wolfers examined thousands of foul calls by NBA basketball referees. They found that, even when controlling for many potentially confounding factors, black players were slightly more likely to receive foul calls from white referee crews than from black referee crews. Likewise, white players had more fouls called on them by black referees than by referees of their own race.[92] These observations strongly suggest that race influences the frequency of foul calls. Moreover, the circumstances surrounding foul calls—including the oversight and fairness safeguards in place, the small magnitude of the

---

[91] *See id.* at 821 ("Thus, although education was seen as more important in this decision, participants still selected the male candidate the majority of the time even when he was less educated.").

[92] *See* Joseph Price & Justin Wolfers, *Racial Discrimination Among NBA Referees* 25–27 (Nat'l Bureau Econ. Research, Working Paper No. 13206, 2007), *available at* http://bpp.wharton.upenn.edu/jwolfers/Papers/NBARace.pdf (finding statistically significant evidence of own-race bias among NBA referees).

overall differences (which were detected only over numerous plays) and the split-second speed at which the calls are made—all strongly suggest that the referees are largely unaware that race affects their decisions.

It is important to realize, however, that the design of the analysis does not permit conclusions about which players are being unfairly victimized by discrimination. Specifically, the data does not necessarily indicate that black players are being discriminated against. The study compares the fouls called on players of different races by black and white referees. It does not, however, provide a baseline measure of the "correct" number of fouls for players of either race. Thus the data is equally consistent with several alternative patterns.

One possibility is that referees of both races are pervasively, if subtly, biased, with all referees systematically favoring their own while disfavoring the opposite race. That is, black and white referees are equally likely to deviate from what is "fair" and make the wrong calls for *all* players. An alternative possibility is that referees of one race—white or black—are altogether fair to everyone, while referees of the opposite race are biased towards everyone—they systematically favor their own players and disfavor opposite race players. But on this scenario it is impossible to tell which group of referees is fair or biased. In particular, the data is potentially consistent with white referees being entirely "fair" to players of both races while black referees are discriminating—against white players and in favor of black—or vice versa. That players of one race receive more foul calls from opposite-race referees is not necessarily inconsistent with this: it could well be that white or black players actually commit more fouls on average or play more aggressively.[93] Thus, although bias of some kind seems to be operating, it is impossible to know whether it runs against or in favor of minority players—or both. Likewise, it is impossible to know which referees' calls are distorted.[94] Because most referees are currently white, the overall effect at present is that white players are subject to relatively fewer foul calls than black players. But, as already noted, this is not necessarily an unfair situation. That is, it is not necessarily the product of an inaccurate bias *against* black players. Likewise, it is impossible to tell whether increasing the number of black referees would increase or decrease biased—as opposed to "correct"—foul calls. Because it is at least possible that only black referees are biased, increasing the number of black referees will not necessarily make the game more fair; it might have the opposite effect. Likewise, adding more white

[93] In fact, the data shows that black players have slightly more fouls per game, but white players (who play fewer minutes per game) have more fouls per minute played. *See id.* at tbl.2. However, these racial differences disappear once a few observable characteristics such as height, weight, and position are controlled. *See id.* at 6 n.3.

[94] *See id.* at 26 ("[U]nfortunately our framework is not well-suited to sorting out whether these results are driven by the actions of black or white referees.").

referees might make the game more, or less, fair; or it might just shift the balance of favoritism from blacks to whites.

In a similar study, Christopher Parsons and his colleagues determined that, after controlling for potentially confounding factors, baseball umpires are more likely to call strikes for pitchers of the same race or ethnicity.[95] That is, umpires seem to favor pitchers from their own group. As with the Price and Wolfers observations, the effects are small and can only be detected when very large numbers of games are evaluated. Similarly, these authors' comparative data do not reveal the direction of the bias relative to any objective baseline of "fair" calls. It is thus impossible to determine whether minorities are the recipients of affirmative action from their same-race umpires, the victims of discrimination from other-race umpires, or both. Likewise, the results do not reveal whether umpires from some groups are fair—and others biased—or whether all umpires are biased. Once again, however, the predominance of white umpires suggests that white players will do better than players from other groups. But whether that advantage is fair or unfair—whether it is the product of unjust discrimination against minorities or the relative paucity of unjust discrimination in *favor* of minorities—is impossible to discern.

Nonetheless, these two studies represent some of the best evidence so far that inadvertent, "subtle" race-based biases might have some influence on patterns of real-world social decision-making. At most, however, they suggest that identity biases might affect some decisions some of the time. And those effects appear quite small—they are detectable only with sophisticated statistical analyses on large data sets involving many decision points. Finally, in keeping with the Norton studies detailed above, it is not possible to predict ahead of time the direction of these biases. That is, on the current state of knowledge it cannot be determined with any confidence whether cognitive biases will favor or disfavor minority group members or whites, or neither, or both. Generalizations are unwarranted: each instance of social decision-making must be evaluated case by case.

Given these vicissitudes, scientists have tried to develop more powerful tools for predicting unconscious discriminatory motive. As already noted, cognitive psychologists have developed a laboratory test of split-second association—the Implicit Association Test (IAT)—that is thought to implicate unconscious processes of stereotyping about groups.[96] The test measures people's tendency to associate negative words with black faces. The significance of that association to real world decision-

---

[95] *See* Christopher A. Parsons et al., *Strike Three: Umpires' Demand for Discrimination* 24 (Nat'l Bureau Econ. Research, Working Paper No. 13665, 2007), *available at* http://www.nber.org/papers/w13665.pdf (finding that umpires express preferences toward players of the same race).

[96] For discussion of the IAT, see *supra* notes 8–10 and accompanying text.

making, however, remains unclear.[97] Yet that has not stopped legal scholars and other commentators from jumping to conclusions. The law review literature is now replete with questionable extrapolations from IAT results to claims of pervasive discrimination against blacks and women throughout social life.[98]

In fact, the evidence that IAT scores correlate or predict real-world discrimination is remarkably thin. On gender, one paper in the literature purports to find an association between high scores on the IAT—indicating a tendency to engage in race and sex based stereotyping—and subjects' willingness to downgrade the social skills of "agentic," competitive, ambitious types of female candidates for feminine-typed, but not masculine-typed, jobs.[99] However, the researchers in that study did not find that the assignment of a lower score to agentic females translated into a lesser willingness to hire them for any job.[100]

For race, the link between the IAT and real-world discrimination has likewise not been established. As stated by Richard Banks and his colleagues, there is thus far "little evidence that Race IAT scores correlate with discrimination against African Americans."[101] One unpublished paper, by T. Andrew Poehlman, collects studies that purport to correlate IAT measures with observable conduct.[102] However, the behaviors that are the focus of the existing research are far different from the types of real-world decisions by managers, judges, and admissions officers that have measurable consequences for social outcomes. Rather, the studies mostly examine the correlation of IAT scores with such "intermediate" behaviors

---

[97] *See, e.g.*, Arkes & Tetlock, *supra* note 8, at 258 (implicit prejudice based on negative associations can have many different causes and manifestations); Banks et al., *supra* note 12, at 1187–88 (stating that it is difficult to characterize behavior as representative of bias). For a discussion of the work of Hart Blanton and his colleagues, see *infra* pp. 40–41.

[98] *See, e.g.*, Kang, *supra* note 11, at 1514 (suggesting that implicit associations can be assumed to translate into behavior); Jerry Kang & Mahzarin R. Banaji, *Fair Measures: A Behavioral Realist Revision of "Affirmative Action,"* 94 CAL. L. REV. 1063, 1064 (2006) (implicit bias influences behavior); Linda Hamilton Krieger & Susan T. Fiske, *Behavioral Realism in Employment Discrimination Law: Implicit Bias and Disparate Treatment*, 94 CAL. L. REV. 997, 1032–33 (2006) (citing the IAT for the proposition that "stereotypes can function as implicit, associative networks that subconsciously predispose the stereotype holder to perceive, characterize and *behave* toward a stereotyped target in stereotype-consistent ways") (emphasis added).

[99] Laurie A. Rudman & Peter Glick, *Prescriptive Gender Stereotypes and Backlash Toward Agentic Women*, 57 J. SOC. ISSUES 743, 759 (2001); *see also* Kang & Banaji, *supra* note 98, at 1091.

[100] Rudman, *supra* note 99, at 756–57 ("It is noteworthy that the IAT did not correlate significantly with the agentic female's hireability ratings in the feminized or the masculine job conditions."). For a critique of this study, see the work of Hart Blanton et al., discussed *infra* pp. 40–41.

[101] *See* Banks et al., *supra* note 12, at 1187 (adding that "[t]he one study we discuss that attempted to correlate IAT scores with shooting [of crime suspects] did *not* find a statistically significant relationship between the two").

[102] *See* T. Andrew Poehlman et al., *Understanding and Using the Implicit Association Test: III: Meta-analysis of Predictive Validity* 32 (Sept. 16, 2005) (unpublished), *available at* http://faculty.washington.edu/agg/pdf/IAT.Meta-analysis.16Sep05.pdf (finding that IAT measures perform well at predicting unconscious and conscious responses).

as body language, eye contact, verbal communication, and friendly gestures towards blacks. Although easily observed in a laboratory setting, these actions are of uncertain significance. As Poehlman concedes, the existing literature says very little about whether the IAT reliably predicts hiring, job promotion, educational evaluation, police conduct, or judicial sentencing.[103]

Hart Blanton and colleagues recently reviewed the handful of published studies that they identified as making explicit claims about the value of IAT test results for predicting discriminatory behavior towards blacks or women.[104] With the goal of assessing the evidence for these claims, the authors contacted the researchers for each study and asked to see their data. No researchers provided a complete set, and the authors of only two of the papers provided any data at all.[105]

For the two papers for which data sets were actually provided, Blanton and his colleagues found that claims of predictive validity were problematic and rested on remarkably weak evidence. Indeed, only one of the studies—by Zeigert and Hanges—attempted to link IAT results to a type of conduct—personnel decisions in the job setting—with legally cognizable consequences. The authors of that paper purported to show that persons scoring higher on the IAT were more likely to discriminate against minority job candidates in an experimental simulation. However, they observed that result only for the experimental condition in which subjects were expressly instructed by managers to disfavor black candidates. In a race neutral condition, subjects' IAT results did not correlate with a propensity to discriminate. In addition to questioning the significance of a finding of discrimination only on express instructions, Blanton and colleagues faulted the results in that paper for other reasons, including the failure to control properly for differences among job candidates being evaluated and irregularities in the researchers' methods for scoring the IAT. In sum, the Zeigert and Hanges study was judged to be of limited

---

[103] *See id.* at 33 (additional research is needed on IAT measures and specific behavior such as police behavior).

[104] *See* Hart Blanton et al., *Strong Claims and Weak Evidence: Reassessing the Predictive Validity of the Race IAT*, J. APPLIED PSYCHOL. (forthcoming 2008) (on file with Connecticut Law Review). The studies reviewed were: Allen R. McConnell & Jill M. Leibold, *Relations among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes* 37 J. EXPERIMENTAL SOC. PSYCHOL. 435 (2001); Jennifer A. Richeson & J. Nicole Shelton, *Thin Slices of Racial Bias*, 29 J. NONVERBAL BEHAV. 75 (2005); Rudman & Glick, *supra* note 99; J. Nicole Shelton et al., *Ironic Effects of Racial Bias during Interracial Interactions*, 16 PSYCHOL. SCI. 397 (2005); Jonathan Zeigert & Paul Hanges, *Employment Discrimination: The Role of Implicit Attitudes, Motivation, and a Climate for Racial Bias*, 90 J. APPLIED PSYCHOL. 553 (2005).

[105] One researcher (McConnell) provided a full data set for the reported results but did not provide raw IAT data. Another (Zeigert) claimed the loss of raw IAT data but provided transformed IAT data and data for the other variables reported. One researcher (Richeson) declined to provide data for the results reported in one paper; his collaborator denied having the data for results reported in another. Finally, the authors of the last paper (Rudman and Glick) asserted they had lost all the data for the results reported in their publication.

value in establishing the predictive validity of the IAT.

The other study for which Blanton and colleagues obtained data—by McConnell and Leibold—looked at the relationship of subjects' IAT scores to the type of "intermediate" behaviors that carry uncertain significance for decisions that affect social rewards.[106] Conduct towards black and white experimental targets was independently assessed by two researcher-judges. "Discrimination" was defined as "avoidant" behavior towards blacks, such as speaking and smiling less often. In re-examining the data for this paper, Blanton and colleagues found significant flaws in the evidence and study design. Specifically, they found that IAT scores significantly predicted the ratings of only one of the two judges in the study. Even then, the subjects' IAT scores were not correlated significantly with behavior toward the black experimenters. Instead, the results were driven by variations in behavior toward the white "controls." More importantly, however, the overall correlation was significantly dependent on a few statistical outliers, or data points that represented extreme, atypical results. In particular, the behavior of a small number (two) of the experimental subjects masked the fundamental trend—or lack thereof—in the data. When the effects of outliers were eliminated, the correlation between the IAT and behavior disappeared.

This careful reassessment of the IAT research reveals that the IAT's power to predict or reveal actual discrimination is, for now, vastly exaggerated. The scientific literature contains remarkably little evidence for a reliable link between IAT measures and the propensity to discriminate against women or protected minorities in the real world. On the current state of knowledge, IAT results do not support any assertions about people's conscious or unconscious discriminatory behavior.

IV. CONCLUSION

Attempts to blame disparities in social outcomes by race or sex on unconscious bias must be approached with caution in the current climate. Without hard evidence, sweeping and categorical claims of unconscious stereotyping are unwarranted. The extent, direction, magnitude, and even existence of unconsciously motivated behaviors against disfavored groups cannot be assumed. Rather, such assertions must be demonstrated. This necessarily requires the careful and patient accumulation of data, as well as rigorous, comprehensive, and sophisticated examinations of existing facts of social life.

---

[106] *See* McConnell & Leibold, *supra* note 104, at 440. The behaviors assessed included: participants' forward body leaning towards experimenters, the extent to which the subjects' body faced the experimenter, the openness of the participants' arms, the expressiveness of the participants' arms, the distance between the experimental confederates and the subjects' chair at the end of the interaction, smiles, speech errors, speech hesitations, fidgeting, and extemporaneous social comments. *Id.* at 438.

In addition, many of the claims concerning unconscious bias do not properly allege category-based bias at all. Rather, they are directed at the disparate impact, or differential effects, of category-neutral criteria. Such patterns are often grounded in "supply side" differences. Existing disparities by race or sex in average qualifications or human capital development affect behavior. These differences influence different groups' overall ability to meet social requirements and compete for social rewards. Contentions sounding in disparate impact, whether implicating conscious or unconscious decision-making processes, should be subject to the types of scrutiny that have long been applied to such claims: neutral rules should be assessed for relevance, efficiency and social usefulness. And even if such rules reflect sound generalizations, there is no avoiding the need to decide whether the costs of eliminating differential impacts are worth the benefits of greater equality of outcome.

Allegations of disparate treatment, based on the accusation that race or sex plays a causal role in social outcomes, should likewise be scrutinized for potential, alternative "supply side" explanations. In particular, well-developed social scientific data indicate that factors other than discrimination can often explain observed group disparities. These alternative possibilities should never be forgotten when considering unconscious bias. Finally, wariness concerning specific claims of inadvertent, as opposed to deliberate or conscious, discrimination is necessary. Many behaviors attributed to unconscious bias could just as well be explained by old-fashioned "rational" or "statistical" discrimination. Such forms of discrimination are nothing new and research in cognitive psychology sheds little light on how they can best be addressed. On the whole, and contrary to dominant opinion, a central role for unconscious stereotyping in social life has yet to be established. The field is in its infancy and much work remains.