Beyond More Accurate Algorithms: Takeaways from McCleskey Revisited

Ngozi Okidegbe
Boston University

Follow this and additional works at: https://repository.law.umich.edu/mlr

Part of the Criminal Law Commons, Law and Race Commons, and the Legal Writing and Research Commons

Recommended Citation
Available at: https://repository.law.umich.edu/mlr/vol121/iss6/15

https://doi.org/10.36644/mlr.121.6.beyond.more

This Review is brought to you for free and open access by the Michigan Law Review at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Michigan Law Review by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
BEYOND MORE ACCURATE ALGORITHMS:
TAKEAWAYS FROM MCCLESKEY REVISITED

Ngozi Okidegbe*


INTRODUCTION

McCleskey v. Kemp1 operates as a barrier to using the Equal Protection Clause to achieve racial justice in criminal administration.2 By restricting the use of statistical evidence in equal protection challenges, McCleskey stifled the power of the discriminatory intent doctrine to combat the colorblind racism emanating from facially neutral criminal law statutes and governmental actions.3 But what if McCleskey had been decided differently? Given that Washington v. Davis4 held that the challenged law or governmental action had to be

---

* Moorman-Simon Interdisciplinary Career Development Associate Professor of Law and Assistant Professor of Computing and Data Sciences, Boston University. For valuable input, support, and feedback, I am grateful to Philip Brink, Pooja R. Dadhania, James E. Fleming, Alexis Hoag-Fordjour, Nicole McConlogue, Jamelia Morgan, Jessica Silbey, I. India Thusi, Mario Barnes, Bennett Capers, Deborah Hellman, Aziz Huq, Linda C. McClain, Kathryn Miller, and Robert L. Tsai. I am also thankful for the invaluable research support from Sydney Sullivan as well as the excellent editorial assistance provided by the book review editors of the Michigan Law Review, especially Gabe Chess and Elena Meth.

1. 481 U.S. 279 (1987). It is important to note that in response to the McCleskey decision stifling the judicial development of equal protection doctrine, abolitionist-oriented defense lawyers began engaging in different and creative litigation strategies in death-penalty cases. See Robert L. Tsai, After McCleskey, 96 S. CAL. L. REV. (forthcoming 2023) (manuscript at 4–6) (on file with author).


3. Eduardo Bonilla-Silva, RACISM WITHOUT RACISTS 2 (5th ed. 2018) (colorblind racism refers to how “whites rationalize minorities’ contemporary status as the product of market dynamics, naturally occurring phenomena, and blacks’ imputed cultural limitations” despite the existence of systemic discrimination).

4. 426 U.S. 229, 239–41 (1976) (holding that a showing of discriminatory impact alone is insufficient to succeed on an equal protection claim—discriminatory purpose is also required).
“traced to a discriminatory racial purpose,” McCleskey have articulated an approach to equal protection doctrine that would have been capable of addressing the sophisticated and sometimes technologically advanced methods by which racial hierarchy is reinforced and protected in criminal administration today?

It is with this question in mind that I read Professor Mario Barnes’s re-written McCleskey decision, which appears as a chapter in Critical Race Judgments: Rewritten U.S. Court Opinions on Race and the Law, edited by Professors Bennett Capers, Devon W. Carbado, R.A. Lenhardt, and Dean Angela Onwuachi-Willig. Using critical race theory, Professor Barnes shows us a different way forward. Critical race theory is an intellectual movement that provides a lens to study the relationship between law and racism. As Professor Capers explains, its aim is to confront and “transform[] the relationship between law and white supremacy to reshape American jurisprudence in a project of racial emancipation and anti-subordination.”

When critical race theory enters the frame, it brings with it a distinct way of knowing about race, as well as racial discrimination, its effects, and potential avenues for its amelioration. For this reason, Professor Barnes offers us more than just an alternative world where Warren McCleskey prevails. He puts forth a framework that would have equipped courts with a set of interdisciplinary and empirical tools to identify and abolish the power of colorblind ideology to encase racially inequitable systems.

To highlight the importance of Professor Barnes’s contribution, this Review will apply Professor Barnes’s framework to a current racial justice challenge: the use of racially biased risk-assessment algorithms within criminal justice.

---

5. Davis, 426 U.S. at 240 (“[T]he basic equal protection principle that the invidious quality of a law claimed to be racially discriminatory must ultimately be traced to a racially discriminatory purpose.”).

6. Professor of Law, University of California, Irvine School of Law.


8. See generally Jessica M. Eaglin, When Critical Race Theory Enters the Law & Technology Frame, 26 MICH. J. RACE & L. (SPECIAL ISSUE) 151 (2021) (discussing how using critical race theory to critique new technologies enables one to identify and contend with the racial implications of new technologies).

9. It is important to note that Professor Barnes’s framework is specifically informed by e-CRT, an empirically grounded form of critical race theory. See, e.g., Mario L. Barnes, Empirical Methods and Critical Race Theory: A Discourse on Possibilities for a Hybrid Methodology, 2016 WIS. L. REV. 443. As Professor Osagie K. Obasogie explains, e-CRT is a movement that pursues “race scholarship in a manner that reflects[] the theoretical orientation put forward by critical race scholarship and also embraces the methodological contributions of social science research.” Osagie K. Obasogie, Foreword: Critical Race Theory and Empirical Methods, 3 U.C. IRVINE L. REV. 183, 185 (2013).
administration. I start by contextualizing how *McCleskey* foreclosed the possibility of using the discriminatory intent doctrine to address the challenge posed by these algorithms. I then introduce Professor Barnes’s framework and imagine how it could be deployed in a current setting. I conclude by addressing implications.

I. CONTEXTUALIZING *MCCLESKEY*

In order to contextualize the impact of *McCleskey*, it is important to briefly lay out terminology. For clarity, I use the term “algorithm” to refer only to risk-assessment algorithms that employ an actuarial method, big data, and information about an individual to produce a forecast about that individual’s future conduct. Jurisdictions are turning to these algorithms in bail, sentencing, and parole as a means to reduce existing racial inequities and reduce incarceration in criminal administration. The optimistic idea is that decisionmakers will rely on this information to make release, detention, sentencing, and parole

10. In applying Professor Barnes’s framework to the problem of racially biased risk-assessment algorithms, the Review necessarily applies the intent prong of the equal protection doctrine (though as modified by Professor Barnes). It should be noted that the intent doctrine has been sharply criticized. Professor Aziz Huq has argued that the very concept of intent (explicit or implicit) is poorly suited to contending with how algorithmic technologies interact with structural inequalities to reproduce social stratification. On this basis, he has advocated for a complete rethinking of equal protection doctrine as it pertains to algorithmic technologies. Aziz Z. Huq, *Constitutional Rights in the Machine-Learning State*, 105 CORNELL L. REV. 1043, 1922–23 (2020). In contending that risk-assessment algorithms reinforce a particular understanding of racism that naturalizes it, Professor Jessica Eaglin’s work implicitly suggests that a focus on intent may be the wrong inquiry for addressing the racial implications of these technologies. Jessica Eaglin, *Racializing Algorithms*, 111 CALIF. L. REV. (forthcoming 2023) (manuscript at 43–49) (on file with author). While I am very sympathetic to this view, expanding the concept of intent along the lines advocated by Professor Barnes does offer one useful starting point (albeit incomplete) for contending with the racial effects of these technologies.

11. It should be noted that the *McCleskey* judgment (as well as Professor Barnes’s alternative judgment) concerned capital cases. Given this, there is an argument that *McCleskey* should not impede a litigant from challenging risk-assessment algorithms under the Equal Protection Clause. However, many scholars have taken the position that *McCleskey* is an impediment (a position that I share). For instance, see Michael Brenner et al., *Constitutional Dimension of Predictive Algorithms in Criminal Justice*, 55 HARV. C.R.–C.L. L. REV. 267, 291 (2020) (“Current equal protection jurisprudence is ill-equipped to address the discrimination brought about by risk-assessment technology. The Supreme Court’s equal protection decisions in *Washington v. Davis* and *McCleskey v. Kemp* appear to foreclose the argument that the use of risk-assessment technology may violate the Equal Protection Clause, as the doctrine stands today.”).


13. See Dorothy E. Roberts, *Digitizing the Carceral State*, 132 HARV. L. REV. 1695, 1718 (2019) (reviewing VIRGINIA EUBANKS, *AUTOMATING INEQUALITY* (2018)); CHRISTOPHER SLOBOGIN, *JUST ALGORITHMS* at vii–viii (2021) (“[T]hese algorithms help judges figure out whether arrested individuals should be released pending trial and whether convicted offenders should receive prison time or an enhanced sentence; they assist parole boards in determining whether to release a prisoner; and they aid correctional officials in deciding how offenders should be handled in prison.”).
decisions without resorting to the racial heuristics that have fueled mass incarceration.  

Even though these algorithms do not explicitly factor in race or use race as an input, they produce racially biased outcomes. I use the term “racially biased algorithm” in two senses. First, I use the term to refer to algorithmic systems that produce inaccurate and inflated predictions of riskiness in regard to racially marginalized individuals as compared to non-racially marginalized individuals. Second, I use the term to refer to algorithmic systems that produce predictions that justify and support the continuation of racial stratification in criminal administration and beyond. Though these two problems tend to intersect, they are distinct. Both senses of “racially biased algorithm” operate in tandem to continue the concentration of carceral control and its physical, psychological, and socioeconomic consequences on people from racially marginalized communities, particularly poor Black communities.

McCleskey is a substantial impediment to challenging racially biased algorithmic systems under current discriminatory intent doctrine. One reason why is that McCleskey restricts the use of statistical evidence, which is typically the only available evidence to show discriminatory intent in the algorithmic context. Another is that, in affirming the judgment, Justice Powell wrote that


15. I use the term “accuracy” to refer to whether the tool reliably predicts the likelihood of the misconduct that it was designed to predict. This issue is generally referred to in the computer literature as the “validity” of the tool, but I use the term “accuracy” since this choice aligns with the common use of the word. It is worth noting that given the current state of equal protection doctrine, another approach to redressing the problem of racial inequality for false positives is to apply the leveling-down doctrine, an approach that Professor Huq discusses outside of the algorithmic context. See generally Aziz Z. Huq, The Discrete Charm of Leveling Down, 90 GEO. WASH. L. REV. 1487, 1527–29 (2022). I would like to thank Professor Huq for bringing this to my attention.


17. Roberts, supra note 13, at 1697.


19. Under McCleskey, a viable equal protection claim can only be made out if an algorithm was designed by a malicious developer who deliberately constructs the algorithm to produce racially biased outcomes. Yet as Professor Huq notes, this does not occur in practice and, even so, “[a]ny moderately competent municipality found using flawed data would hardly concede that it was doing so intentionally.” Aziz Z. Huq, Racial Equity in Algorithmic Criminal Justice, 68 DUKE L.J. 1043, 1093 (2019).

20. McCleskey is not the only impediment obstructing meaningful change in this regard. As Professor Huq has explained, the requirement of discriminatory intent put forth by the Supreme Court in Washington v. Davis remains ill-defined in the case law and presents conceptual difficulties when applied to the algorithmic context. Id. at 1085, 1088–94 (discussing Washington v. Davis, 426 U.S. 229 (1976)).
the discriminatory intent standard could not be satisfied by evidence demonstrating that a facially neutral criminal statute produced racial disparities within a state system that itself had a long history of explicit racial discrimination. Rather, a claimant must show that a jurisdiction "enacted or maintained [the challenged law] because of an anticipated racially discriminatory effect." Otherwise, courts will assume that there is a legitimate and nondiscriminatory explanation for the racial disparities identified. This requirement is insurmountable for a claimant in the algorithmic context for two reasons. First, adopting jurisdictions can claim to rely on these algorithms to redress racial inequities in spite of their racial effects as opposed to because of them. Second, those jurisdictions can always point to a nondiscriminatory, yet circular, explanation as to why an algorithm predicted a racially marginalized person as at high risk for future misconduct: the individual has traits that are considered risk factors by the algorithm.

II. MCCCLESKEY REREWRITTEN

This brings us to MCCCleskey rewritten. Writing as a justice, Professor Barnes takes a radically different approach to the discriminatory intent doctrine. First, he rejects a narrow interpretation of discriminatory intent for the reason that the doctrine must be able to account for "the complex ways in which racism operates" (p. 560). This purposive approach to discriminatory intent frees the doctrine from its current preoccupation with conscious discrimination, allowing the doctrine to take account of unconscious discrimination, systemic discrimination, and future ways in which racial hierarchy might manifest itself.

Second, Professor Barnes’s framework takes seriously valuable insights from interdisciplinary scholarship and empirical work. Unlike Justice Powell, Professor Barnes encourages courts to accept empirical social-science evidence on the view that such evidence is vital for comprehending how race structures the design, construction, and implementation of the system, institution, or law challenged. The takeaway is that legal professionals must look

22. Id. Note that on this point the Court was reaffirming and quoting its prior decision in Pers. Admin. v. Ffeeney, 442 U.S. 256, 279 (1979) ("Discriminatory purpose] implies that the decisionmaker . . . selected or reaffirmed a particular course of action at least in part 'because of,' not merely 'in spite of,' its adverse effects upon an identifiable group.").
23. MCCCCleskey, 481 U.S. at 298.
24. Huq, supra note 19, at 1093 (explaining that, on account of Ffeeney, a jurisdiction can escape liability for the racial effects of an algorithmic technology—particularly if those effects are caused by the use of flawed data—by contending that it had not intentionally selected the algorithm because of its racial effects, but instead in spite of them).
27. See pp. 561–64.
28. See p. 563.
to other fields to develop the multidimensional competency needed to grapple with racism. As Professor Barnes puts sharply, “[t]o ignore the insights of [social science] fields of expertise is to ignorantly attempt to divorce the law from the social context in which it was created and in which it functions, in itself an act of willful blindness” (p. 563). Moreover, under Professor Barnes’s framework, statistical evidence of discrimination is admissible, probative, and sufficient on its own to prove discriminatory intent. This allowance stems from his recognition that statistical evidence may be the sole available evidence of discrimination, and requiring a claimant to forgo it in such cases “would be to impose a burden so heavy it would be impossible to bring a discrimination claim”—a warning that materialized for McCleskey himself and resulted in his execution four years later.

Finally, and importantly, Professor Barnes takes seriously the country’s history of racism and the gravity of racism’s operation in criminal administration. For this reason, his framework requires courts tasked with adjudicating an equal protection claim to factor in a jurisdiction’s history of explicit racial discrimination and the stakes of the law, policy, or action at issue. Moreover, a finding that a state has a history of explicit racial discrimination and has a current racial disparity (as demonstrated by validated statistical evidence) gives rise to the presumption that, as Professor Barnes proffers, “purposeful discrimination, albeit potentially operating at an unconscious level,” exists (p. 576). To justify this facially race-neutral but presumably race-conscious system, the burden would shift to the state to “provide a compelling interest to justify continuing its current race-conscious system, and why this manner of considering race is the least restrictive means of achieving the state’s goals.” In other words, the presence of historical and explicit racial discrimination combined with a current racial disparity subjects a jurisdiction to strict scrutiny on the view that the jurisdiction is operating a prima facie racist system.

29. See p. 568.
30. See p. 570.
32. See p. 561 (including “relevant documented legal and social history and social science research in the jurisdiction” as well as “consideration of the gravity of the law being applied” as part of the totality-of-the-circumstances analysis for discriminatory intent); p. 571 (“[I]ntent is broad enough a concept that it can be satisfied, at times, when a state has a history of racial discrimination coupled with current evidence of systemic but uncorrected racial disparity.”).
33. See p. 576.
34. It is important to note that Barnes takes the position that if a jurisdiction is confronted with a history of explicit discrimination as well as statistical evidence of current racial disparity and does not intervene to eliminate these disparities, then it will be presumed that the state had selected this biased course of action because of (and not in spite of) its adverse effects upon the identifiable group. The Barnes presumption is in tension with the Feeney decision, as Barnes notes: “One difference between our reasoning in this case and the analysis in Feeney is our approach here presumes that legislatures who are confronted with such data and do nothing are de
Applying Professor Barnes’s approach to the problem of racially biased algorithms demonstrates the importance of his contribution. As an initial matter, Professor Barnes’s approach to the doctrine of discriminatory intent would enable effective equal protection challenges to colorblind algorithms producing racially biased predictions. One reason is the permittance of statistical evidence as proof, which would provide a path for claimants to meet their evidentiary burden under the discriminatory intent doctrine. Another reason is Professor Barnes’s approach to strict scrutiny, which would be applicable in this context.

Present-day algorithms have their origins in the risk-assessment instruments developed in the early twentieth century. These instruments were used in parole determinations and embodied racist tropes around criminality. For this reason, as Professor Bernard Harcourt reminds us, “[t]hroughout most of the twentieth century, race was used [within these instruments] explicitly and directly as a predictor of dangerousness.” Consequently, risk-assessment instruments operated historically to justify racial subordination and containment under the veneer of scientific objectivity. While today’s algorithms do not use race as an input, they do reproduce racial disparities. This historical discrimination and contemporary context would trigger strict scrutiny under Professor Barnes’s framework and place a burden that most jurisdictions (if not all) would be unable to meet. There is no compelling governmental purpose achieved by continuing to use algorithms that produce inaccurate and inflated predictions about racially marginalized individuals’ future misconduct. Even if a compelling purpose did exist (for instance, a jurisdiction might state that the purpose is to protect public safety), the use of inaccurate algorithms would not be the least restrictive means to achieve the jurisdiction’s aim. For the above reasons, the application of Professor Barnes’s approach would provide claimants with injunctive, declaratory, or even monetary relief. Moreover, it could mean setting aside an order denying pretrial release, entrance into an alternative-to-incarceration program, or parole if that denial was conditioned upon a racially biased algorithmic prediction. This would

38. See Roberts, supra note 13, at 1710 (noting how “removing human discretion from sentencing only compounded racial disparities in the criminal justice system”).

facto choosing a biased system because it works in this manner, rather than in spite of that fact.”
provide a constitutional path toward the use of more accurate algorithms at the very least.39

But Professor Barnes’s rewritten judgment has broader implications, too. As Professor Benjamin Eidelson has argued, accurate algorithms would sustain racial stratification.40 This is because the factors that elevate a person’s risk of engaging in violent conduct—such as poverty, pollution, unclean water, housing, and employment instability—are the very conditions that Black and other politically oppressed communities are disproportionately forced to live in.41 As Professor Allegra McLeod contends, this state-sanctioned vulnerability and economic deprivation is the source of violent and otherwise harmful conduct.42 Thus, building accurate algorithms means building algorithms that will still produce racially disparate predictions. Moreover, as a result, accurate algorithms would continue to add to or, as Professor Deborah Hellman has theorized, “compound” the injustices already experienced by racially marginalized communities in this society, rather than reduce these injustices.43 Given this, accurate algorithmic predictions in this racially stratified world would only serve to reproduce this racially stratified world, though perhaps on a smaller scale.

Nevertheless, accuracy as a goal dominates algorithmic-fairness literature. This is in part due to the fact that accurate algorithmic predictions would lead to the release of a significant number of Black and other racially marginalized people who are currently detained. In fact, many have argued that current algorithmic predictions are preferable to judicial assessments of risk, given that the biases that are held by and influence decisionmakers are harder to identify.

39. It is important to note that when I use the term “more accurate algorithms” here I am referring to the fact that Barnes’s framework would enable a claimant to effectively challenge the fact that the algorithm in use leads to racial inequality in its production of false positives. The term is not meant to suggest that Barnes’s framework would lead to more accurate algorithms as a general matter. This distinction is important because the problem of accuracy (outside of the context of racial inequality) is a matter of due process and not a matter within the purview of equal protection doctrine. A consideration of the relationship between accuracy norms and due process is beyond the scope of this Review. For a consideration of this point, please see generally Huq, supra note 10.


42. See id.

43. Professor Deborah Hellman puts forth a theory regarding the moral objection of using big data and algorithmic decisionmaking. The concern relates to the fact that these algorithmic technologies use inputs that are often associated with an individual’s experience of injustice. Examples include an individual’s experience with poverty, poor health, child abuse, and job instability. When a person (who has experienced socially created injustice, such as child abuse) is subjected to an algorithm that uses that injustice as a risk factor, that algorithm compounds the prior injustice experienced by that person. For more information on this point, see Deborah Hellman, Big Data and Compounding Injustice, J. MORAL PHIL. (forthcoming 2023) (manuscript at 4, 14–16) (on file with author).
and challenge than the biases affecting algorithms. It is here that lies the power of Professor Barnes’s rewritten decision. It allows us to extricate ourselves from this dichotomy where we are forced to choose between racially inequitable algorithms or the racially inequitable use of judicial discretion, since both would be unconstitutional under Professor Barnes’s approach.

Instead, we can find a path forward beyond what is understood in this current moment to be “an improvement over the status quo,” a benchmark that is often employed and necessarily raises the “for whom” question. By engaging with liberatory imaginaries such as critical race theory, we can envision a world in which accurate algorithms constitute an equal protection violation for facilitating the continuation of a racially stratified system of criminal administration.

CONCLUSION

In a world that takes critical race theory seriously, the algorithms of today have no place. But the technology that produced them might. Figuring out if and on what terms algorithms might remain in use would require having a difficult conversation about what racial justice within criminal administration means. For guidance, we need only to turn to Derrick Bell’s reflection on Brown v. Board of Education, where he emphasized the importance of centering impacted communities—a reflection relied upon by Professor Barnes himself in drafting his rewritten decision. Addressing the multifaceted nature of racism within criminal administration requires engaging with a diversity of stakeholders, with particular attention to the communities that stand to be the most impacted by any technological reforms. It means accounting for the knowledge produced by those communities whose expertise is discredited and ignored in the technological realm. It also means ceding decisionmaking power over if and how any technology might be used to achieve racial justice.

44. Michael Selmi, Algorithms, Discrimination and the Law, 82 OHIO ST. L.J. 611, 632 (2021) (“If it appears that the algorithm is producing discriminatory results, it can be altered to address that discrimination. This may not always be successful but between correcting a discriminatory algorithm and correcting human biases, again, the smart money should be on the algorithm.”); Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan & Cass R. Sunstein, Discrimination in the Age of Algorithms, 10 J. LEGAL ANALYSIS 1, 1–2 (2018) (describing how “[h]uman decisions are frequently opaque to outsiders, and they may not be much more transparent to insiders” and arguing that “when algorithms are involved, proving discrimination will be easier—or at least it should be, and can be made to be”).


47. See pp. 561–63.

48. See Ngozi Okidegbe, Discredited Data, 107 CORNELL L. REV. 2007, 2046–58 (2022) (explaining that solving the problem of algorithmic discrimination involves shifting toward non-carceral knowledge sources, including “the community knowledge sources relied upon by the communities most affected by the criminal legal system”).
to the communities democratically excluded from the paradigm governing these technologies in our present day.49

Achieving this feat might be impossible. This is particularly so since the impact of actualizing Professor Barnes’s framework would extend beyond the technological context to the entirety of the criminal system. Taken to its natural conclusion, it would require the invalidation of most (if not all) criminal law statutes, interventions, and decisions, given their racial, classed, and otherwise socially inequitable underpinnings and consequences. This country may not be willing or ready to fully grapple with the permanence of racism50 and the deep work required to make a world more equitable than the one before us. But doing this work is a precondition to unlocking a world in which our technological advancements, alongside our legal systems, honor the spirit of the Fourteenth Amendment, and a world where racial justice activists have the tools needed to keep it that way.

49. See Ngozi Okidegbe, The Democratizing Potential of Algorithms?, 53 CONN. L. REV. 739, 767–77 (2022) (discussing the importance of shifting power over algorithmic governance to “members from the most impacted communities”).