Stanford Law Review Online

Volume 79 September 2026

ESSAY

Antidiscrimination Law in an AI-Pervaded World

Ralph Richard Banks, Guy-Uriel Charles, Spencer Overton & Victor Y. Wu*

Introduction

In March 2026, Stanford Law School hosted a two-day conference on artificial intelligence (AI) and antidiscrimination law, organized by law faculty at George Washington, Harvard, and Stanford.[1] It brought together two groups that rarely engage with each other: AI researchers and technology company representatives on one side, and civil rights lawyers, nonprofit leaders, and law professors on the other.[2] The convening was closed to allow candid discussion—no recording, no attribution, and no public identification of participants. This Essay takes up questions raised by the convening and argues that AI may reinforce racial and other group inequalities in ways current legal doctrine is poorly equipped to handle.

The law's ability to address discriminatory AI faces both immediate and long-term challenges. In the short term, the Trump Administration has expressed hostility to the disparate impact approach—a doctrine of particular

* Ralph Richard Banks is the Jackson Eli Reynolds Professor of Law at Stanford Law School and the Faculty Director of the Stanford Center for Racial Justice. Guy-Uriel Charles is the Charles J. Ogletree Jr. Professor of Law and the Faculty Director of the Charles Hamilton Houston Institute for Race and Justice at Harvard Law School. Spencer Overton is the Patricia Roberts Harris Research Professor of Law and the Faculty Director of the Multiracial Democracy Project at George Washington University Law School. Victor Y. Wu is a Ph.D. Candidate in Political Science at Stanford University and a Data Science Scholar at the Stanford Institute for Human-Centered AI. He received his J.D. from Stanford Law School in 2025. This Essay has benefited enormously from comments provided by Jennifer Eberhardt. Thanks to Hannah L. Dahleen of the *Stanford Law Review Online* for her insights and editorial assistance.

1. *See Conference Addresses Impact of AI on Antidiscrimination Law*, STAN. REP. (Mar. 19, 2026), https://perma.cc/XUV4-BVZU.

2. *See Antidiscrimination Law x AI Convening*, STAN. L. SCH., https://perma.cc/9M56-A7H8 (archived Aug. 27, 2026).

importance in the age of AI.[3] President Trump issued an executive order condemning disparate impact and directing agencies to "eliminate the use of disparate impact liability in all contexts to the maximum degree possible," to reconsider pending proceedings and consent decrees relying on the doctrine, and to explore federal preemption of state-level disparate impact protections.[4] In December 2025, the Department of Justice finalized a rule rescinding the disparate impact provision of its Title VI regulations.[5]

This Essay focuses on a longer-term threat to antidiscrimination law, one that will persist after President Trump leaves office, no matter who succeeds him: the current Supreme Court's interpretation of the Equal Protection Clause. The Court's construal of the clause to mandate colorblindness may have far-reaching implications,[6] possibly invalidating laws that require or even simply permit a response to AI's production of disparate group outcomes. What is at stake is the equality ethos that has long animated antidiscrimination law. Giving substance to that value will require bolstering the disparate impact approach for the algorithmic age.[7]

Part I of this Essay explains why AI is such a consequential development for governmental and institutional decisionmaking and why it will often recreate or perpetuate racial disparities, even in the absence of any intent to

3. *See* Ralph Richard Banks & Victor Y. Wu, *From Intent to Impact: Equal Protection in the Age of Artificial Intelligence*, 2026 U. ILL. L. REV. (forthcoming 2026) (manuscript at 127-30) (on file with authors).

4. Exec. Order No. 14281, 90 Fed. Reg. 17537 (Apr. 28, 2025).

5. *Rescinding Portions of Department of Justice Title VI Regulations to Conform More Closely with the Statutory Text and to Implement Executive Order 14281*, 90 Fed. Reg. 57141 (Dec. 10, 2025); *see also* Alex Guillén & Hassan Ali Kanu, *DOJ Rolls Back Anti-Discrimination Rules*, POLITICO (Dec. 9, 2025, 3:58 PM EST), https://perma.cc/T6WR-2VQ4. The Administration also moved to erase the data needed to detect and remedy discrimination, by, for example, repealing guidance encouraging schools to collect and analyze racial data on student discipline and to consider whether disparities might indicate unlawful discrimination. Exec. Order No. 14280, 90 Fed. Reg. 17533 (Apr. 28, 2025). Most recently, in June 2026, the Office of Legal Counsel issued a pair of opinions extending the attack on disparate impact: one concluding that the Equal Employment Opportunity Commission's guidance implementing Title VII's disparate impact provisions is unconstitutional, and a second concluding that the Equal Credit Opportunity Act does not create disparate impact liability at all. *See* Constitutionality of Disparate-Impact Liability Under Title VII, 50 Op. O.L.C. (June 9, 2026) (slip op. at 1-2), https://perma.cc/7BFV-B7G9; Whether the Equal Credit Opportunity Act Creates Disparate Impact Liability, 50 Op. O.L.C. (June 12, 2026) (slip op. at 1), https://perma.cc/KBX9-CC3J ("The Equal Credit Opportunity Act does not create disparate-impact liability.").

6. *See, e.g.*, Ralph Richard Banks & Victor Y. Wu, *Racial Entanglement and the Colorblindness Fallacy*, 2026 UTAH L. REV. (forthcoming 2026) (manuscript at 8-25) (on file with authors).

7. *See* Banks & Wu, *supra* note 3.

discriminate and without relying on racial categories. Part II identifies the Court's increasing embrace of colorblindness as a threat even to voluntary efforts to assure the fairness of AI with respect to race. Part III briefly sketches two additional considerations for antidiscrimination law in an AI-pervaded world.

I. Pervasive AI and Persistent Disparity

A. The Ubiquity of AI Decisionmaking

Artificial intelligence is poised to reshape large parts of American life, causing some to hope for the best and others to fear the worst.[8] Government at every level has already adopted AI,[9] and that reliance will only deepen as the technology matures and as agencies find new uses for it. The private sector is likewise rapidly incorporating AI into its business processes. Job applications are being screened by an algorithm before any human assesses a candidate's qualifications.[10] Mortgage and other credit decisions often rely on AI,[11] as do admissions and other decisions made by colleges and universities.[12]

In sum, AI likely will become ubiquitous in most facets of life, including policing, education, housing, credit markets, benefits determinations, and more.[13] The implications of these developments are less that AI will wholly

8. *Compare* Michael R. Strain, *The Case for AI Optimism*, 60 NAT'L AFF., Summer 2024, at 65, 65-66, 78, 82-83, https://perma.cc/X5NV-Y8E8, *with* Kristin Burnham, *These Are the Most Urgent AI Risks, According to 272 Experts*, MIT SLOAN (July 20, 2026), https://perma.cc/WKL2-3S3D.

9. *See* SANAM HOOSHIDARY, CHELSEA CANADA & WILLIAM CLARK, NAT'L CONF. STATE LEGISLATURES, ARTIFICIAL INTELLIGENCE IN GOVERNMENT: THE FEDERAL AND STATE LANDSCAPE 1 (2024), https://perma.cc/MA2F-2R2R; DAVID FREEMAN ENGSTROM, DANIEL E. HO, CATHERINE M. SHARKEY & MARIANO-FLORENTINO CUÉLLAR, GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 6 (2020), https://perma.cc/ANS5-JL2Y.

10. Danielle Abril, *AI Is Now Screening Job Candidates Before Humans Ever See Them*, WASH. POST (July 1, 2025), https://perma.cc/D7JB-PFWF.

11. *See* Robert Bartlett, Adair Morse, Richard Stanton & Nancy Wallace, *Consumer-Lending Discrimination in the FinTech Era*, 143 J. FIN. ECON. 30, 40-41 (2022).

12. *See, e.g.*, Alice Scott, *UNC Uses AI in Admissions Review Process, Documents Show How*, THE DAILY TAR HEEL (Jan. 20, 2025), https://perma.cc/FSN8-ARJH.

13. *See generally* VIRGINIA EUBANKS, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR (2018) (documenting automated eligibility, risk scoring, and triage systems targeting poor families); CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY (2016) (examining how opaque predictive models in education, credit, and policing entrench inequality); RUHA BENJAMIN, RACE AFTER TECHNOLOGY: ABOLITIONIST TOOLS FOR THE NEW JIM CODE (2019) (arguing that ostensibly neutral technologies encode racial hierarchy in a "New Jim Code").

displace humans than that AI and human decisionmaking will become so intertwined that the line between the two will be difficult, if not impossible, to draw.[14]

B. Racial Impact Without Racial Categories

While AI systems lack any intent to discriminate,[15] they often produce substantial and troubling racial disparities.[16] Racially disparate outcomes may result from the phenomenon that scholars have termed "proxy discrimination"—the tendency of an AI system to functionally recreate racial categories through reliance on informationally relevant characteristics and markers that are associated with race.[17] Even now, high-profile workplace-discrimination lawsuits are challenging the use of AI by employers.[18]

Consider, for instance, a lender that uses a model to predict default. Race is not among the model's inputs. But the model learns that applicants who shop at certain retailers, hold accounts at certain banks, or live within certain census tracts are more likely to default. These variables are predictive in part because of what they reveal about race: residential segregation, disparities in inherited wealth, and unequal access to mainstream credit all mean that racial position is embedded in ordinary consumer behavior. The model is doing exactly what it was built to do, and in doing so, it produces racial disparities.

The massive datasets on which AI models are trained will encompass innumerable variables linked to race—including those linked subtly, in ways that exceed our human awareness or ability to catalogue. Race is such a salient identity and category in American society precisely because it is associated with myriad sociological, economic, political, and psychological factors, many of which are markers of inequality.[19] It is difficult, in fact, to imagine an area of

14. *Cf.* Rebecca Crootof, Margot E. Kaminski & W. Nicholson Price II, *Humans in the Loop*, 76 VAND. L. REV. 429, 432 (2023) ("Artificially intelligent algorithms are being integrated into decisionmaking processes at mind-boggling speed and scale.").

15. Chiraag Bains, *The Legal Doctrine That Will Be Key to Preventing AI Discrimination*, BROOKINGS (Sept. 13, 2024), https://perma.cc/AKN2-EUP8 ("Intent is an expression of will that we ascribe to human beings. Inanimate machines are not capable of intent. They simply do what their designers and operators instruct them to do, even if the instruction is not deterministic.").

16. *See, e.g.*, Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 673-74 (2016).

17. Michael Carl Tschantz, *What Is Proxy Discrimination?*, 2022 ACM CONF. FAIRNESS, ACCOUNTABILITY & TRANSPARENCY 1993, 1993-94.

18. *See, e.g.*, Mobley v. Workday, Inc., 740 F. Supp. 3d 796, 801-02 (N.D. Cal. 2024); Complaint ¶¶ 1, 3-4, Does 1 Through 26 v. Meta Platforms, Inc., No. 26-cv-07122 (N.D. Cal. July 13, 2026), ECF No. 1.

19. *See, e.g.*, MICHAEL OMI & HOWARD WINANT, RACIAL FORMATION IN THE UNITED STATES 105-06, 109-12, 124-27 (2014); Maya Sen & Omar Wasow, *Race as a Bundle of Sticks:*
footnote continued on next page

American life free from racial meanings or associations.[20] Race is linked not only to where people live, work, and attend school; it is linked also to their cultural orientation and personal preferences and behavior—everything from how people walk and talk to their taste in art and music.[21] And as AI models are trained on ever larger and more varied data (not simply written words, but images, audio, and video), the risk of proxy discrimination will only grow.

The use of racial proxies in decisionmaking models may be difficult to avoid, just as is the models' tendency to functionally recreate race while being blind to the formal category. When a model is told to identify the factors that predict an outcome, in a society where race is entangled with residence, wealth, schooling, and health—current conditions themselves shaped in part by a history of racially discriminatory law and public policy[22]—the factors that predict the relevant outcome will often be associated with race. The result is a system blind to race as a formal category yet one which reliably reproduces its consequences, thus generating disparities traceable neither to discriminatory purpose nor to racial classification.[23]

Three concepts should be distinguished: (i) an AI system may produce racially disparate outcomes; (ii) it may produce those outcomes by relying on variables that function as proxies for race; and (iii) a decisionmaker may attend to such disparities—e.g., by selecting among models on that basis—in order to reduce them. The first two involve no consideration of race by any human decisionmaker. Only the third does, and that, as Part II explains, is what the Court's colorblindness jurisprudence imperils.

Designs that Estimate Effects of Seemingly Immutable Characteristics, 19 ANN. REV. POL. SCI. 499, 510 (2016).

20. *See generally* JENNIFER L. EBERHARDT, BIASED: UNCOVERING THE HIDDEN PREJUDICE THAT SHAPES WHAT WE SEE, THINK, AND DO (2019) (drawing on social science research to show the pervasiveness of unconscious racial bias).

21. *See id.* at 14.

22. Rashida Richardson, *Racial Segregation and the Data-Driven Society: How Our Failure to Reckon with Root Causes Perpetuates Separate and Unequal Realities*, 36 BERKELEY TECH. L.J. 1051, 1070-71 (2021) (explaining that "racial segregation inevitably influences and shapes data sources, the data mining processes, and human biases and practices in the technology development process"); *see also* RICHARD ROTHSTEIN, THE COLOR OF LAW: A FORGOTTEN HISTORY OF HOW OUR GOVERNMENT SEGREGATED AMERICA 13-14 (2017) (describing how "federal, state, and local governments purposely created segregation in every metropolitan area of the nation").

23. The opacity of AI compounds the problem. *See* Banks & Wu, *supra* note 3. As AI systems become more complex, developers, deployers, regulators, and affected individuals may have difficulty predicting, identifying, or explaining the disparities they produce. *Id.*

II. Disparate Impact and the Colorblind Constitution

A. An Uncertain Doctrine

The racial disparities produced by AI implicate the disparate impact framework, as they would neither be amenable to discriminatory-purpose analysis nor rely on formal racial classification. On the former, the discriminatory intent standard becomes unworkable, if not incoherent, as applied to the development or deployment of an AI system. The challenges of proving discriminatory intent on the part of the developer or deployer are daunting, given the opacity of increasingly complex AI systems; such complex systems resist explanation in terms that would support an inference about a human's intent.[24] On the latter, the functional recreation of race by AI would likely not count as racial classification due to the stringent and formal manner in which the Supreme Court has defined that concept.[25]

While the Supreme Court has rejected disparate impact claims under the Equal Protection Clause,[26] disparate impact claims have been incorporated into each of the three landmark federal statutes from the Civil Rights era: (i) the Civil Rights Act of 1964, which prohibited discrimination across various settings, including public accommodations, employment, and programs receiving federal funds,[27] (ii) the Voting Rights Act of 1965,[28] and (iii) the Fair Housing Act of 1968.[29]

---

24. *See, e.g.*, Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J.L. & TECH. 889, 897-905 (2018); Vikas Hassija et al., *Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence*, 16 COGNITIVE COMPUTATION 45, 53-54 (2024).

25. *See, e.g.*, Hernandez v. New York, 500 U.S. 352, 361-62 (1991) (plurality opinion) (holding that a criterion that "might well result in the disproportionate removal of prospective Latino jurors" was nonetheless race neutral, because "[e]qual protection analysis turns on the intended consequences of government classifications" and disparate impact alone "does not turn the prosecutor's actions into a *per se* violation of the Equal Protection Clause"); *see also id.* at 375 (O'Connor, J., concurring in the judgment) ("No matter how closely tied or significantly correlated to race the explanation for a peremptory strike may be, the strike does not implicate the Equal Protection Clause unless it is based on race.").

26. *See* Washington v. Davis, 426 U.S. 229, 238-39 (1976).

27. Civil Rights Act of 1964, Pub. L. No. 88-352, 78 Stat. 241 (codified as amended in scattered sections of 42 U.S.C.).

28. Voting Rights Act of 1965, Pub. L. No. 89-110, 79 Stat. 437 (codified as amended at 52 U.S.C. §§ 10101, 10301-702).

29. Fair Housing Act of 1968, Pub. L. No. 90-284, tit. VIII, 82 Stat. 73, 81-89 (codified as amended at 42 U.S.C. §§ 3601-31).

The text of these laws clearly prohibited intentional discrimination,[30] but the status of the disparate impact approach was always a bit uncertain. For example, disparate impact claims became part of employment law not through the text of the statute, but instead because of a Supreme Court decision. In 1971, the Court held in *Griggs v. Duke Power Company* that Title VII of the 1964 Civil Rights Act prohibited facially neutral employment practices (such as the standardized tests and high school diploma requirements at issue in that case) if they disproportionately excluded racial minority groups without sufficient justification.[31]

After the Supreme Court in *Wards Cove* weakened the disparate impact claim,[32] Congress codified disparate impact in the 1991 Amendments to the Civil Rights Act.[33] Despite the paucity of current disparate impact cases,[34] the doctrine remains alive. In the Fair Housing Act (FHA), disparate impact claims result from judicial interpretations of statutory language, to which federal agencies and Congress acquiesced. Lower federal courts recognized disparate impact claims for decades, and the Supreme Court in 2015 confirmed in *Texas Department of Housing and Community Affairs v. Inclusive Communities Project* that the FHA permits challenges to neutral policies that disproportionately harm protected groups, provided plaintiffs can trace the disparity to a specific policy and the defendant cannot show a valid justification.[35]

The Voting Rights Act, by contrast, came to provide for disparate impact claims in its statutory text. Section 2 as enacted in 1965 closely tracked the language of the Fifteenth Amendment, and in *City of Mobile v. Bolden* (1980) a plurality read the statute, consistent with the Amendment, to reach only voting practices motivated by a discriminatory purpose.[36] Congress responded in 1982 by amending Section 2 to establish a "results" test, under which plaintiffs in a vote-dilution case could prevail on a showing of discriminatory effect, without proof of discriminatory purpose.[37]

---

30. *See* sources cited *supra* notes 27-29.

31. *See* Griggs v. Duke Power Co., 401 U.S. 424, 431-32, 436 (1971).

32. Wards Cove Packing Co. v. Atonio, 490 U.S. 642, 656-57 (1989).

33. Civil Rights Act of 1991, Pub. L. No. 102-166, 105 Stat. 1071 (codified as amended in scattered sections of the U.S. Code).

34. Michael Selmi, *Was the Disparate Impact Theory a Mistake?*, 53 UCLA L. REV. 701, 763 (2006).

35. Tex. Dep't of Hous. & Cmty. Affs. v. Inclusive Cmtys. Project, Inc., 576 U.S. 519, 534-35, 542 (2015).

36. 446 U.S. 55, 60-65 (1980) (plurality opinion).

37. Voting Rights Act Amendments of 1982, Pub. L. No. 97-205, § 3, 96 Stat. 131, 134 (codified as amended at 52 U.S.C. § 10301); *see* Thornburg v. Gingles, 478 U.S. 30, 35 (1986) (describing Congress's response). To be clear, discriminatory effect is not all plaintiffs need to show to prevail in a vote-dilution claim. *See Gingles*, 478 U.S. at 50-51.

footnote continued on next page

In the last quarter century, however, the Court has become increasingly hostile to statutory disparate impact claims. *Alexander v. Sandoval* eliminated the private right of action for disparate impact claims under Title VI of the 1964 Civil Rights Act (applicable to any entity receiving federal funding), leaving agency enforcement as the only means of enforcing such claims.[38]

In sum, disparate impact supplies the principal framework capable of reaching AI-produced disparities,[39] as such disparities can arise without discriminatory purpose or formal racial classification. Yet the doctrine rests on statutory rather than constitutional foundations—and even those statutory foundations are rickety. And the doctrine now faces a fundamental threat: the Supreme Court interpreting the Constitution to forbid the attention to group outcomes on which disparate impact necessarily depends.

B. Colorblindness Ascendant

The Supreme Court's embrace of the constitutional ideal of colorblindness imperils the disparate impact approach. As Chief Justice Roberts memorably declared in the *Parents Involved* case concerning voluntary school integration: "The way to stop discrimination on the basis of race is to stop discriminating on the basis of race."[40] The linchpin of that perspective is the rejection of the benign/invidious asymmetry associated with the anti-subordination conception of equal protection. The anti-subordination approach condemns race-based practices that aggravate patterns of group hierarchy, but not race-based practices that aim to promote integration or undermine historical patterns of group inequality.[41] The Court has jettisoned that asymmetry in a series of cases, beginning with *Bakke*,[42] and extending through *Croson*[43] and *Adarand*.[44] In *Students for Fair Admissions v. Harvard*,[45] the Court applied strict scrutiny to end race-conscious admissions in both public and private

And as we discuss below, the Court has since substantially narrowed the results test. *See infra* notes 47-50 and accompanying text.

38. 532 U.S. 275, 293 (2001).

39. Banks & Wu, *supra* note 3.

40. Parents Involved in Cmty. Schs. v. Seattle Sch. Dist. No. 1, 551 U.S. 701, 748 (2007).

41. On the anti-subordination conception, a racial classification meant to dismantle a racial hierarchy is not the constitutional equivalent of one that is meant to maintain it, and the two therefore should be viewed differently. The colorblind perspective collapses that distinction and treats every governmental use of race as equally suspect regardless of whether it is meant to dismantle or maintain a racial hierarchy.

42. Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265, 294-95 (1978) (opinion of Powell, J.).

43. City of Richmond v. J.A. Croson Co., 488 U.S. 469, 493-94 (1989) (plurality opinion).

44. Adarand Constructors, Inc. v. Pena, 515 U.S. 200, 223-24, 227 (1995).

45. Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2175 (2023).

universities, resting on the unchallenged presumption that conduct violating the Equal Protection Clause likewise violates Title VI.[46]

Most recently, the Court decided *Louisiana v. Callais* in April 2026.[47] That case concerned a challenge by non-Black voters to Louisiana's decision to add a second majority-Black district to its congressional map, after lower courts concluded that a map with only one majority-Black district likely diluted the Black vote in violation of Section 2 of the Voting Rights Act. In holding that the alleged vote dilution did not warrant the creation of an additional majority-minority district, the Court substantially narrowed the circumstances in which Section 2 calls for a race-conscious districting remedy, emphasizing that the totality-of-circumstances inquiry should focus on evidence bearing on present-day intentional racial discrimination in voting.[48] As Justice Kagan argued in dissent, the Court's approach moves the Section 2 inquiry substantially toward discriminatory intent, notwithstanding Congress's 1982 adoption of an effects test.[49] *Callais* thus increased the stringency of the colorblindness mandate, in effect requiring plaintiffs to come awfully close to proving intentional discrimination to warrant a race-based response. Five weeks later, in an unsigned order applying *Callais* on the emergency docket, a majority of the Court referred approvingly, for the first time ever, to "our colorblind Constitution."[50]

To be clear, *Callais* arose in the redistricting context, and does not by its terms disturb the disparate impact provisions of Title VII or the Fair Housing Act. But its reasoning need not be confined to districting. Section 2's effects test is perhaps the clearest instruction Congress has ever given decisionmakers to attend to racial effects. If even Section 2's congressionally dictated effects test is inadequate to justify race-conscious action, it is not obvious why the more

46. *Id.* at 2156 n.2 (noting that the Court has treated discrimination violating the Equal Protection Clause by a federally funded institution as a Title VI violation, then declining to reconsider that premise because no party had asked it to do so). The Court did reserve the question of the constitutionality of affirmative action at the nation's service academies, which the Court said raised special considerations. *Id.* at 2166 n.4.

47. Louisiana v. Callais, 146 S. Ct. 1131 (2026).

48. *Id.* at 1160 (concluding that "the 'totality of circumstances' inquiry must focus on evidence that has more than a remote bearing on what the Fifteenth Amendment prohibits: present-day intentional racial discrimination regarding voting").

49. *Id.* at 1176 (Kagan, J., dissenting) ("At each of *Gingles*'s steps, the majority imposes new proof requirements, serving a common objective: to convert an effects test, as commanded by Congress, into a purpose test, as preferred by this Court.").

50. Allen v. Milligan, 146 S. Ct. 1377, 1380 (2026) (per curiam) ("In *Louisiana v. Callais* . . . to resolve the tension between vote-dilution claims under § 2 of the Voting Rights Act of 1965 and our colorblind Constitution, we updated the standards for § 2 liability established by *Thornburg v. Gingles* . . . .").

equivocal foundations of the disparate impact approach in other settings would fare better.

Indeed, the Court's decision in *Callais* highlights a broader issue raised in the Court's 2009 decision in *Ricci v. DeStefano*.[51] In *Ricci*, the city of New Haven administered a civil service exam for promotions within the fire department, and then, after white test takers as a group scored higher than their Black counterparts, declined to make any promotions, opting instead to develop a different exam, which presumably would have less disparate impact on Black firefighters. A group of white firefighters and one Hispanic firefighter challenged the city's decision as racially disparate treatment in violation of Title VII of the 1964 Civil Rights Act. The Court sided with the plaintiffs, discounting the city's claim that its decision was necessary to avoid liability for disparate impact.

In a concurrence, Justice Scalia emphasized the constitutional precariousness of statutory disparate impact law. He wrote that "Title VII's disparate impact provisions place a racial thumb on the scales, often requiring employers to evaluate the racial outcomes of their policies, and to make decisions based on (because of) those racial outcomes. That type of racial decisionmaking is . . . discriminatory."[52] Justice Scalia referred to a coming "war" between disparate impact statutes and equal protection, and it was clear how he thought it should end.[53]

Callais has moved us closer to that war. Or rather, *Callais* might signal how most of the current Justices want that war to end. The mandate of constitutional colorblindness could be leveraged to invalidate the disparate impact provisions of statutory law.[54] The logic would be the same that informed the application of the racial classification rule to invalidate affirmative action and other policies enacted to promote racial inclusion or integration: the rejection of the benign/invidious asymmetry. On that view, discarding the test results because the high scorers were white would be no more permissible than doing so because the high scorers were Black.[55]

---

51. Ricci v. DeStefano, 557 U.S. 557 (2009); *see also* Richard Primus, *Equal Protection and Disparate Impact: Round Three*, 117 HARV. L. REV. 493, 494-96 (2003) (asking whether equal protection forbids statutory disparate impact standards and concluding that it does not).

52. *Ricci*, 557 U.S. at 594 (Scalia, J., concurring).

53. *Id.* at 595.

54. To avoid the constitutional problem, the Court might instead do in other areas of law what it did in *Callais*: reinterpret a disparate impact provision to in fact require intentional discrimination. Either way, the idea of disparate impact as a harm that might warrant racial awareness would disappear.

55. Indeed, under this approach, a goal of racial integration itself might be viewed as a discriminatory purpose. *See* Michelle Adams, *Is Integration a Discriminatory Purpose?*, 96 IOWA L. REV. 837, 839-41 (2011). While the Court has not so held, certain Justices have

footnote continued on next page

Just weeks after *Callais*, the Office of Legal Counsel (OLC) concluded that the Equal Employment Opportunity Commission's guidance implementing Title VII's disparate impact provisions is unconstitutional, on the ground that the guidance pressures employers into race-conscious decisionmaking.[56] The opinion recasts disparate impact liability solely as an evidentiary device for detecting intentional discrimination.[57] An OLC opinion is not binding on courts, and it cannot displace *Griggs* or the 1991 Amendments to the Civil Rights Act. But it illustrates the reach of the Court's reasoning: the war Justice Scalia anticipated is now being waged in the executive branch on the basis of a voting rights decision.

The application of such reasoning to AI-produced racial disparities should be clear. Suppose an employer tests a screening model, finds that it selects white applicants at a substantially higher rate than Black applicants, and—in the hope of narrowing that disparity—decides to retrain the algorithm using reweighted data or to adopt a different model that reduces both accuracy and the race gap. Both of these responses could be viewed, consistent with the reasoning of *Ricci*, as racially discriminatory.[58] The logic of *Callais* is that only an effort to counter intentional discrimination would warrant taking race into account.

---

signaled a willingness to move in this direction. *See* Coal. for TJ v. Fairfax Cnty. Sch. Bd., 146 S. Ct. 541, 544-45 (2024) (Alito, J., dissenting from denial of certiorari) (arguing that a facially race-neutral admissions policy may nonetheless constitute intentional race discrimination when adopted with the purpose of altering a school's racial composition).

56. Constitutionality of Disparate-Impact Liability Under Title VII, *supra* note 5, at 1 ("EEOC's Title VII guidelines are unconstitutional because they contemplate liability based on disparate effects alone, without regard to an employer's likely intent, and pressure employers to engage in race-based decisionmaking.").

57. *See id.* at 2 ("Rather than treating disparate impact as an evidentiary mechanism to smoke out intentional discrimination—imposing liability only when disproportionate adverse effects give rise to a strong inference of intentional discrimination—EEOC's historic interpretations contemplate liability based on disproportionately adverse effects alone, without regard to an employer's likely intent."); *see also id.* at 1 ("Properly understood, disparate-impact liability proscribes only those practices that reflect a significant likelihood of intentional discrimination.").

58. *See* Daniel E. Ho & Alice Xiang, *Affirmative Algorithms: The Legal Grounds for Fairness as Awareness*, 2020 U. CHI. L. REV. ONLINE 134, 137 (2020) (explaining "why modern antidiscrimination law is likely to consider most algorithmic fairness methods to be forms of affirmative action"). *But see* Pauline Kim, *Race-Aware Algorithms: Fairness, Nondiscrimination and Affirmative Action*, 110 CALIF. L. REV. 1539 (2022) (arguing that race-aware debiasing of AI systems does not involve racial classification and therefore should not trigger strict scrutiny). It is important to note that *Ricci* might be interpreted narrowly in light of its circumstances (in particular, that the city did not announce its intention to discard the exam scores until after the administration of the test, for which many test takers had spent months preparing).

Our point is descriptive, not justificatory. The Equal Protection Clause need not be (and in our view should not be) interpreted to disable efforts to identify and remedy racial inequality. But the Court's recent reasoning may be extended precisely in that direction, as demonstrated by the recent OLC opinion.[59]

III. Two Further Challenges for Antidiscrimination Law

Reconstructing antidiscrimination law in the age of AI will require more than simply applying existing disparate impact law, whose viability is necessary but not sufficient for antidiscrimination law to grapple with the challenges posed by AI. AI warrants a fuller and more robust analysis of outcomes than is customary under existing disparate impact doctrine. The pervasive use and advance of AI will require rethinking two aspects of disparate impact law.

A. Accuracy and Fairness

While this Essay has primarily considered disparate impact through an anti-subordination lens (according primacy to the harm of racially disparate outcomes), in practice, the doctrine accords as much or more weight to the legitimacy of a defendant's justification for the challenged practice. An admittedly simplified distillation of the law is that sufficiently significant racial disparities shift the burden to the defendant to demonstrate a sufficiently important justification for the challenged policy. If the defendant meets that burden, fairness concerns yield, at that step, to a legitimate institutional interest. In the final phase of the framework, however, the plaintiff can still prevail by identifying a less discriminatory alternative that the defendant refuses to adopt.

Challenges to AI-inflected decisionmaking will place unusual pressure on this framework. Predictive accuracy (a presumably legitimate and weighty interest) will be proffered as a justification for the challenged practice. And at the final stage, the plaintiff's effort to identify a less discriminatory alternative will confront a circumstance that is not unique to AI but that will be accentuated and made more salient by it: when groups are in fact dissimilarly situated, predictively accurate AI will nonetheless generate group disparities that cannot all be eliminated.[60] More precisely, the elimination or diminution

59. *See supra* notes 56-57 and accompanying text.

60. *See* Sandra G. Mayson, *Bias In, Bias Out*, 128 YALE L.J. 2218, 2249 (2019) ("When base rates of the predicted outcome differ across groups, the most accurate algorithm possible will predict that outcome at different rates across groups. Imposing certain metrics of output equality will therefore have a cost in accuracy.").

of one disparity will often necessarily give rise to another. And efforts to lessen disparities might compromise accuracy.[61]

A model might be evaluated against any of several measures of group fairness: whether it selects members of each group at similar rates, whether its Type I or Type II errors fall on each group at similar rates, and whether a given score means the same thing for each group. A model that satisfies one of these measures typically cannot satisfy another. There will rarely be a single, unitary "adverse impact." There will instead be a splintering of impacts, requiring some assessment of the relative costs (and benefits) of the different disparities, and a weighing of the disparities against the institutional interest in accuracy. This is the sort of calculus for which disparate impact doctrine provides little guidance. The challenge is that AI will require methods to evaluate multiple, competing forms of disparity and to navigate increasingly difficult tradeoffs among fairness, accuracy, and institutional objectives.

Conceptually, AI flips a common intuition about the relation between bias and rationality. During the civil rights era that birthed contemporary antidiscrimination law, barriers to opportunity generated both error and disparities. Eliminating such barriers through disparate impact law would often further both accuracy and fairness. The use of AI models, as they become more sophisticated and rely on ever larger and more varied bodies of information,[62] may set those goals against each other. To the extent the models pick up on inequalities that both predict relevant outcomes and are associated with race, disparities may increase even as AI becomes more accurate.[63] AI may

61. Such tradeoffs have been extensively examined in the algorithmic fairness literature. On the incompatibility of competing fairness measures, see Alexandra Chouldechova, *Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments*, 5 BIG DATA 153, 157 (2017). On the accuracy costs of imposing fairness constraints, see, for example, Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel & Aziz Huq, *Algorithmic Decision Making and the Cost of Fairness*, 23 PROCS. ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING 797, 797 (2017). *But see* Kit T. Rodolfa, Hemank Lamba & Rayid Ghani, *Empirical Observation of Negligible Fairness—Accuracy Trade-offs in Machine Learning for Public Policy*, 3 NATURE MACH. INTEL. 896, 896 (2021) ("Our results suggest that trade-offs between fairness and effectiveness can in fact be negligible in practice . . . ."). For a treatment situating these results within antidiscrimination law, see generally Deborah Hellman, *Measuring Algorithmic Fairness*, 106 VA. L. REV. 811, 829-31 (2020) (exploring "the epistemic and practical significance of both accuracy and the type of error (false positive or false negative) in individual cases").

62. *See, e.g.*, Veronika Samborska, *Scaling Up: How Increasing Inputs Has Made Artificial Intelligence More Capable*, OUR WORLD IN DATA (Jan. 20, 2025), https://perma.cc/67AK-HDFU ("[B]igger AI systems, trained on more data and using more computational resources, tend to perform better.").

63. *See* Aleksandar Popovic & Martin R. Huecker, *Study Bias*, STATPEARLS (2023), https://perma.cc/QT8N-KSEZ ("Bias is colloquially defined as any tendency that limits impartial consideration of a question or issue."). Of course, this is not always the case.

footnote continued on next page

therefore reproduce structural inequality, even as the effects-based legal framework best suited to identifying and evaluating that inequality becomes increasingly precarious.

B. Algorithmically Constructed Groups

Antidiscrimination law has long rested on the idea of group categorization. Suspect classifications such as race or gender are accorded special scrutiny.[64] But AI calls into question the contours of the categories on which antidiscrimination law operates. Disparate impacts in the age of AI may no longer neatly run along lines of race or non-race. Rather, models that sort people along hundreds or thousands of dimensions at once will produce groupings that no one specified in advance and that correspond to no category the law recognizes. As Janneke Gerards and Frederik Zuiderveen Borgesius explain: "[A]lgorithms can generate new categories of people based on seemingly innocuous characteristics, such as web browser preference or apartment number, or more complicated categories combining many data points."[65] Some of these groupings will functionally recreate race. Others will resemble race without tracking it.

Ultimately, AI-driven disparities will burden groups defined through an agglomeration of features in our social world. Such groups may in some ways resemble racial groups and in other ways may not.[66] The world we are entering may be one where the very groups themselves will be in part the product of the data collected on individuals and the ways that AI operates. That is a hard problem for a body of law organized around a fixed list of protected classes. Race and sex existed as social categories long before the law took notice of them. New groups produced by AI will come into being only as the models

---

Skewed outcomes might also result from incomplete or misleading data inputs. *See, e.g.*, Julie Rogers & Alexandra Jonker, *What Is Data Bias?*, IBM (Oct. 4, 2024), https://perma.cc/4CN4-JRQX.

64. Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll., 143 S. Ct. 2141, 2220 (2023) (Gorsuch, J., concurring) ("[C]ourts apply strict scrutiny for classifications based on race, color, and national origin; intermediate scrutiny for classifications based on sex; and rational-basis review for classifications based on more prosaic grounds.").

65. Janneke Gerards & Frederik Zuiderveen Borgesius, *Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence*, 20 COLO. TECH. L.J. 1, 1 (2022).

66. *See* Anita L. Allen, *Dismantling the "Black Opticon": Privacy, Race, Equity, and Online Data-Protection Reform*, 131 YALE L.J. F. 907, 923-26 (2022) (describing "racialized sorting by humans and machines" and explaining how characteristics including names, residential addresses, and associations can operate as racial proxies); *id.* at 932-33 (calling for regulation of automated "profiling" and "identification" and discrimination based on race or "characteristics that are its proxies").

that generate them are deployed. They may not be identifiable until after the harm is done. The emergence of such groups does not make traditional protected classes obsolete or render every machine-generated grouping legally or normatively equivalent to historically subordinated groups. The challenge will be determining when, if ever, novel algorithmic groupings warrant legal concern without flattening the distinctive history and status of established protected classes.

Conclusion

AI may well eventually make or shape decisions in nearly every area of life that antidiscrimination law was written to govern: hiring, lending, housing, schooling, medical care, policing, and more. In doing so, it will reduce racial disparities in some areas and produce them in others. When AI produces disparities, it will likely do so without anyone intending to discriminate. Neither of the two doctrinal categories the Supreme Court has built its equal protection jurisprudence around—discriminatory intent and racial classification—will have much to say about this. Disparate impact is the principal framework for reaching AI-driven disparities that arise without discriminatory intent or formal racial classification.

That framework is weak, and it is getting weaker. It has no constitutional foundation in the eyes of the current Court, and its statutory footings are a patchwork, assembled over decades by different means in different areas of law. The current administration is dismantling what it can reach. And the Court's colorblindness cases suggest that it might finish the job. The Court's colorblindness mandate not only threatens to eviscerate enforcement of the disparate impact approach; it also threatens the ability of institutions to police themselves.

The constitutional danger we identify is not a reason to abandon disparate impact, but instead a reason to defend the legitimacy of attending to unequal effects even when those effects arise without discriminatory intent or formal classification. Antidiscrimination law should be reconstructed by building on, rather than abandoning, the effects-based principles embodied in disparate impact.

AI presents new challenges for disparate impact. The doctrine assumes that disparate outcomes signal something gone wrong. But persistent racial disparities are perfectly consistent with the increasing accuracy of AI models. A new decisional calculus will be necessary, one that does not presuppose that accuracy will lessen disparities. The disparate impact framework also assumes that the burdened group is one the law already recognizes; in fact, AI will sort people into groups that no legislature has ever named. The emergence of such groups need not diminish the distinct legal and normative significance of

historically protected classes, but it will require antidiscrimination law to revisit some of its assumptions.

The convening that motivated this Essay brought together people who rarely sit in the same room. That was the point: to join the expertise of technologists with that of legal scholars and lawyers. The equality ethos that has animated antidiscrimination law for sixty years will survive the coming decades only if it is rebuilt to identify and evaluate AI-driven disparate impacts that may arrive without intent, without classification, and increasingly without visibility.