Maryland Law Review

Volume 80 | Issue 3 Article 1

The Intersection of Race and Algorithmic Tools in the Criminal Legal System

Vincent M. Southerland

Follow this and additional works at: https://digitalcommons.law.umaryland.edu/mlr

Part of the Law and Race Commons

Recommended Citation
Vincent M. Southerland, *The Intersection of Race and Algorithmic Tools in the Criminal Legal System*, 80 Md. L. Rev. 487 (2021)
Available at: https://digitalcommons.law.umaryland.edu/mlr/vol80/iss3/1

This Article is brought to you for free and open access by the Academic Journals at DigitalCommons@UM Carey Law. It has been accepted for inclusion in Maryland Law Review by an authorized editor of DigitalCommons@UM Carey Law. For more information, please contact smccarty@law.umaryland.edu.

THE INTERSECTION OF RACE AND ALGORITHMIC TOOLS IN THE CRIMINAL LEGAL SYSTEM

VINCENT M. SOUTHERLAND*

A growing portion of the American public—including policymakers, advocates, and institutional stakeholders—have accepted the fact that racism endemic to the United States infects every stage of the criminal legal system. Acceptance of this fact has resulted in efforts to address and remedy pervasive and readily observable systemic bias. Chief among those efforts is a turn toward technology—specifically algorithmic decision-making and actuarial tools. Many have welcomed the embrace of technology, confident that technological tools can solve a problem—race-based inequity—that has bedeviled humans for generations. This Article engages that embrace by probing the adoption of technological tools at various sites throughout the criminal legal system and exploring their efficacy as a remedy to racial inequality. Then, by applying a racial justice lens, this Article develops and offers a set of prescriptions designed to address the design, implementation, and oversight of algorithmic tools in spaces where the promise offered by technological tools has not been met. Adherence to that lens may draw us closer to what this Article terms a pragmatic abolitionist ethos regarding the use of technological tools in the criminal legal system. Such an ethos does not mean the immediate absence of a criminal legal system altogether. It instead means a criminal system that ultimately operates in ways dramatically different from the current regime by divesting from incarceration and investing in community well-being, human welfare, and rehabilitation.

©2021 Vincent M. Southerland.

*Assistant Professor of Clinical Law and Co-Faculty Director, Center on Race, Inequality, and the Law at New York University School of Law. I thank Deborah Archer, Rachel Barkow, Nina Chernoff, Andrew Ferguson, Randy Hertz, Sandra Mayson, Erin Murphy, Melissa Murray, Ngozi Okidegbe, Ndidi Oriji, Rashida Richardson, Jason Schultz, Maneka Sinha, Kim Taylor-Thompson, and Anthony Thompson for their invaluable insights and helpful comments on earlier drafts and Steven Demarest, Samuel Dunkle, Chiraayu Gosrani, Zachary Mason, Jeremy Burton, Kameron Johnston, and Kimberly Fayette for their excellent research assistance. This Article also benefitted from thoughtful feedback provided by the attendees of the 10th Annual John Mercer Langston Writing Workshop, including Mario Barnes, Paul Butler, I. Bennett Capers, Sean Hill, and Matthew Shaw. I am especially grateful to the editors of the *Maryland Law Review* for their tremendous editorial assistance and incredibly hard work throughout the publication process. All errors are my own.

INTRODUCTION

> Bubbling crack, jewel theft and robbery to combat poverty
> And end up in the global jail economy
> Stiffer stipulations attached to each sentence
> Budget cutbacks but increased police presence
> And even if you get out of prison still living
> Join the other five million under state supervision
> This is business: no faces, just lines and statistics
> From your phone, your Zip Code to SSI digits
> The system break man, child, and women into figures
> Two columns for "who is" and "who ain't [n*****]"
> Numbers is hard and real and they never have feelings
> But you push too hard, even numbers got limits
> – Mos Def[1]

> [T]he great force of history comes from the fact that we carry it within us, are unconsciously controlled by it in many ways, and history is literally present in all that we do.
> – James Baldwin[2]

* * *

No matter where one stands on matters of law and order, the problems that characterize America's criminal legal system are well-documented. It is rife with inequity and plagued by unfairness. More often than not, criminal legal system outcomes turn on the characteristics, identity, and economic status of those targeted by the system and the actors responsible for its operation. It is overly punitive, generally devoid of empathy, and, in large part, fails to ensure public safety, individual accountability, or the health of communities.[3]

---

1. MOS DEF, *Mathematics*, *on* BLACK ON BOTH SIDES (Rawkus Records 1999).
2. James Baldwin, *The White Man's Guilt*, EBONY, Aug. 1965, at 47 (emphasis omitted).
3. *See generally* MICHELLE ALEXANDER, THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS (2010); RACHEL BARKOW, PRISONERS OF POLITICS: BREAKING THE CYCLE OF MASS INCARCERATION (2019); PAUL BUTLER, CHOKEHOLD: POLICING BLACK MEN (2018); DANIELLE SERED, UNTIL WE RECKON: VIOLENCE, MASS INCARCERATION, AND A ROAD TO REPAIR (2019); *see also* Mark Osler, *Short of the Mountaintop: Race Neutrality, Criminal Law, and the Jericho Road Ahead*, 49 U. MEM. L. REV. 77, 87–90 (2018) (describing racial inequality that permeates each stage of the criminal legal system); Radley Balko, *There's Overwhelming Evidence the Criminal-Justice System Is Racist. Here's the Proof*, WASH. POST, (June 10, 2020) https://www.washingtonpost.com/graphics/2020/opinions/systemic-racism-police-evidence-

Critiques of the criminal legal system that are rooted in race and racism have exposed it as a mechanism of social control, designed from birth to perpetuate an oppressive regime of racial caste and fueled by an irrational fear of people of color.[4] History has witnessed the system evolve over time—in part through intentional design, in part through benign neglect, and in part through policies that ignore structural inequality to exacerbate harm—to consume communities of color and relegate its subjects to second-class citizenship.[5] These critiques have laid bare the deeply problematic decision-making that characterizes the criminal system.

In recent years, those concerned about the failings of the system, including policymakers, data scientists, technologists, and system actors have turned to technology as a means of curing its ills.[6] They have done so with good intentions. Many want to eliminate racism in the system and implement policies in service of that goal, such as shrinking the prison and jail populations, ending money bail, and holding system actors to account for bias.[7] Data-driven, fact-based, technological interventions that inform the decision-making of system actors are thought of as the solution.

---

criminal-justice-system/ (detailing a myriad of studies demonstrating racial inequity in the criminal legal system).

4. Bryan Stevenson, *Slavery Gave America a Fear of Black People and a Taste for Punishment. Both Still Define our Criminal-Justice System*, N.Y. TIMES MAG. (Aug. 14, 2019), https://www.nytimes.com/interactive/2019/08/14/magazine/prison-industrial-complex-slavery-racism.html; *see also* Loïc Wacquant, *Deadly Symbiosis: When Ghetto and Prison Meet and Mesh*, 3 PUNISHMENT & SOC'Y 95, 97 (2001) (viewing the criminal legal system, and the "penal system as an instrument for the management of dispossessed and dishonored groups," such that "the astounding upsurge in black incarceration in the past three decades as a result of the obsolescence of the ghetto as a device for caste control and the correlative need for a substitute apparatus for keeping (unskilled) African Americans . . . in a subordinate and confined position in physical, social, and symbolic space"); *see also* PAUL BUTLER, CHOKEHOLD: POLICING BLACK MEN 5 (2017) (observing that "the system is broke on purpose").

5. Barack Obama, *The President's Role in Advancing Criminal Justice Reform*, 130 HARV. L. REV. 811, 816–20 (2017); *see also supra* note 4.

6. The hope that technology might resolve America's racial ills is longstanding. As early as 1967, civil rights leader Roy Wilkins asked whether the computer could "turn its impersonal, unprejudiced magic upon our agonizing race problem? Could it not, after digesting the facts which whites and blacks have fogged over for so long give us an outline of our obligation? [C]an [] not the computer become a guidepost to interracial justice and peace?" CHARLTON MCILWAIN, BLACK SOFTWARE 243 (2020).

7. *See* Sam Corbett-Davies, Sharad Goel, & Sandra González-Bailón, *Even Imperfect Algorithms Can Improve the Criminal Justice System,* N.Y. TIMES (Dec. 20, 2017), https://www.nytimes.com/2017/12/20/upshot/algorithms-bail-criminal-justice-system.html (extolling the potential value of technology to address inequity, bias, racism and other harmful facets of the criminal legal system); Stephanie Wykstra, *Philosopher's Corner: What is "Fair"? Algorithms in Criminal Justice*, ISSUES IN SCIENCE AND TECHNOLOGY 34, no. 3 (Spring 2018), https://issues.org/perspective-philosophers-corner-what-is-fair-algorithms-in-criminal-justice/ (same); Alex P. Miller, *Want Less-Biased Decisions? Use Algorithms.*, HARV. BUS. REV. (July, 26, 2018), https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms (same).

Unfortunately, as presently envisioned and executed, the turn to technological tools is destined to fall short of its lofty and admirable aims. This Article suggests that to transform the criminal legal system, advocates need to adopt a lens centered on racial justice to inform technology-based efforts rather than simply layering tools onto it in its current state. Doing so would mean that rather than attempting to solve or eradicate racism, we would account for the role that racism plays as we design, implement, and engage in oversight of these technological tools. This Article applies a racial justice-focused theoretical framework grounded in critical race theory to confront and address the pressing problems presented by the use of technological tools in the criminal legal system. Ultimately, this framework suggests that we deploy such tools in a way that represents a paradigmatic shift in the way our current criminal system operates.

The problem technology purports to solve is not new.[8] Decision-making in the criminal legal system depends on assessments by human beings at every critical stage of the criminal process, from arrest and prosecution to punishment and reintegration. The decisions that people make about other people—and the social and cultural baggage attached to those decisions—are among the clearest sources of systemic inequity. While some attempt to turn a blind eye to this reality, the quantitative data and qualitative experience make clear that the discretion of criminal legal system actors is infused with bias.[9]

It is within that context, and with a growing acknowledgment of these truths, that efforts to reform the criminal legal system have been undertaken.

8. Before algorithmic tools emerged, assessments of risk were based largely on individual judgements—gut instinct informed by experience. Sarah L. Desmarais & Evan M. Lowder, *Pretrial Risk Assessment Tools: A Primer for Judges, Prosecutors, and Defense Attorneys*, SAFETY AND JUST. CHALLENGE 1, 5 (2019), http://www.safetyandjusticechallenge.org/wp-content/uploads/2019/02/Pretrial-Risk-Assessment-Primer-February-2019.pdf. Bias was the norm, decision-making was wildly inconsistent, and disparities emerged and grew. *Id*. The last decade or so has brought with it the development and use of actuarial tools to help judges and other actors forecast outcomes and make better decisions. *Id*.

9. Mona Lynch & Marisa Omori, *Crack as Proxy: Aggressive Federal Drug Prosecutions and the Production of Black—White Racial Inequality*, 52 LAW & SOC'Y REV. 773, 799–803 (2018) (concluding that prosecutorial discretion was a significant driver of racial disparities in the sentences received for crack related offenses); Osler, *supra* note 3, at 79, 92 (describing "the swamp where [racial] inequality breeds: the hidden world of discretion" and noting that "[t]he criminal justice system as a whole abounds with unobserved discretion that masks the influence of racial bias"); Robert J. Smith & Justin D. Levinson, *The Impact of Implicit Racial Bias on the Exercise of Prosecutorial Discretion*, 35 SEATTLE U. L. REV. 795, 805–22 (2012) (describing the "broad and deep" nature, and potential impact of, implicit racial bias on prosecutorial discretion, and the racially skewed system the exercise of that discretion produces); L. Song Richardson, *Systemic Triage: Implicit Racial Bias in the Criminal Courtroom*, 126 YALE L.J. 862, 887 (2017) ("[T]he enormous discretion wielded by prosecutors, defense lawyers, and judges facilitates racial bias, both conscious and implicit.").

Previous waves of reform have looked to shape or somehow improve decision-making by systemic actors to eradicate the implicit and explicit bias that fosters injustice.[10] The introduction of artificial intelligence, predictive analytics, automated decision-making, actuarial risk assessment instruments, and machine learning—collectively known in this Article under the general umbrella of algorithmic tools—into the criminal legal field is, in some ways, just the latest attempt to improve decision-making and counter the frailties of human judgment.[11]

Proponents of algorithmic tools market them to criminal legal system reformers and stakeholders as a novel approach with greater potential than

10. The federal sentencing guidelines are one example of this kind of change that ultimately failed to yield the fairness reformers hoped they would. *See* Rachel E. Barkow, *Sentencing Guidelines at the Crossroads of Politics and Expertise*, 160 U. PA. L. REV. 1599, 1609, 1619 (2012) (describing the adoption of the sentencing guidelines as "a story about the desire for racial justice. Unfortunately, even a cursory look at criminal justice in the United States—in states with or without guidelines—demonstrates that questions of racial justice have hardly been answered."). The guidelines were intended to weed out disparities and curtail wild inconsistencies in judgments about individuals, while providing common factors that judges could use to consider and determine risk. *Id.* at 1619–22. They were successful to a point, because they did, at times, narrow the previously wide divergences in assessments and imposed greater uniformity in outcomes. *Id*. But those gains in efficiency were outweighed by the loss of a system that imposed individualized justice for those who came before the court. *Id.*

11. An algorithm is an unambiguous set of steps undertaken to solve a problem. *Algorithm*, MERRIAM-WEBSTER DICTIONARY (11th ed. 2020); Jacob Brogan, *What's the Deal With Algorithms?*, SLATE (Feb. 2, 2016, 10:29 AM), https://slate.com/technology/2016/02/whats-the-deal-with-algorithms.html. A helpful and straightforward definition of algorithmic tools are those which apply "an automated protocol to a large volume of data to classify new subjects in terms of the probability of expected criminal activity and in relation to the application of state coercion." Aziz Z. Huq, *Racial Equity in Algorithmic Criminal Justice*, 68 DUKE L.J. 1043, 1060 (2019). This definition consists of an algorithm, or an automated protocol, to "routinize[ ] a decision" about state intrusion into one's life. *Id.* The sheer volume of data examined requires the use of an algorithm and advanced computing resources. *Id.* The tools are also forward looking, in that they make predictions about future behavior or occurrences, rather than identifying historical instances of criminal activity. *Id.* In this context, an algorithm is "any well-defined computational procedure that takes some value, or set of values, as *input* and produces some value, or set of values, as *output*." Gabriel Nicholas, *Explaining Algorithmic Decisions*, 4 GEO. L. TECH. REV. 711, 714 (2020) (emphasis on original) (quoting THOMAS H. CORMEN ET AL., INTRODUCTION TO ALGORITHMS 5 (3rd ed. 2009)). An algorithm can refer to rules that produce a deductive output, or "to inductive procedures that come up with their own rules by generalizing from examples. Algorithms that use the latter option are called *machine learning algorithms*." *Id*. (emphasis in original). Machine learning uses an algorithm to create an algorithm. *Id*. Specifically, it uses one algorithm, a learner, to read data as a set of numerical features, infer rules about those features to predict a particular outcome, and then produce a model that embodies those rules. *Id*. I am using the term algorithmic tools to generally refer to a range of processes and tools, but recognize that there are differences in the methods, techniques, and development of each type of tool. *See* Vincent M. Southerland & Andrea Woods, *What Does Fairness Look Like? Conversations on Race, Risk Assessment Tools, and Pretrial Justice*, CTR. ON RACE, INEQUALITY, AND THE L. AT N.Y.U. LAW & AM. CIV. LIBERTIES UNION, 1, 5–6 (Oct. 2018) (highlighting definitional differences).

all past reform mechanisms.[12] This is not surprising given the great promise that these technologies purport to hold. They have been deployed in an attempt to forecast where crimes may take place,[13] to identify potential perpetrators and crime victims,[14] to predict one's risk of re-arrest or appearance in court,[15] to determine an appropriate sentence,[16] and to suggest when one should be released from incarceration.[17] Proponents of the tools and stakeholders who have embraced them have heralded them as race neutral, countering one of the most persistent and pernicious concerns with the criminal legal system.[18] And they have been posited as improving outcomes for all.[19]

12. Sam Corbett-Davies, *supra* note 7; Anne Milgram, *Why Smart Statistics are the Key to Fighting Crime*, TED (Oct. 2013) (transcript available at https://www.ted.com/talks/anne_milgram_why_smart_statistics_are_the_key_to_fighting_crime/transcript); Adam Neufeld, *In Defense of Risk-Assessment Tools*, MARSHALL PROJECT (Oct. 22, 2017, 10:00 PM), https://www.themarshallproject.org/2017/10/22/in-defense-of-risk-assessment-tools;, *Bail Reform*, ARNOLD VENTURES, https://www.arnoldventures.org/work/release-decision-making/ (last visited Dec. 12, 2019). In part, this turn to technology exemplifies what software developer and data journalist Meredith Broussard has called "technochauvinism," which is "the belief that tech always the solution." MEREDITH BROUSSARD, ARTIFICIAL UNINTELLIGENCE: HOW COMPUTERS MISUNDERSTAND THE WORLD 9 (2018).

13. Randy Rieland, *Artificial Intelligence is Now Used to Predict Crime. But Is It Biased?*, SMITHSONIAN MAG. (Mar. 5, 2018), https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/.

14. Jeff Asher & Rob Arthur, *Inside the Algorithm That Tries to Predict Gun Violence in Chicago*, N.Y. TIMES (June 13, 2017), https://www.nytimes.com/2017/06/13/upshot/what-an-algorithm-reveals-about-life-on-chicagos-high-risk-list.html.

15. *See generally* Robert Werth, *Risk and Punishment: The Recent History and Uncertain Future of Actuarial, Algroithmic, and "Evidence-based" Penal Techniques*, 13 SOCIO. COMPASS 126 (2019).

16. Sara Chodosh, *Courts Use Algorithms to Help Determine Sentencing, but Random People Get the Same Results*, POPULAR SCI. (Jan. 18, 2018), https://www.popsci.com/recidivism-algorithm-random-bias/.

17. Amy McCaig, *Algorithms For Parole Can Have Serious Bias Problems*, FUTURITY (Apr. 9, 2019), https://www.futurity.org/risk-assessment-tools-prison-2031222/.

18. *See, e.g.,* Alex Chohlas-Wood & E. S. Levine, *A Recommendation Engine to Aid in Identifying Crime Patterns*, 49 INFORMS J. ON APPLIED ANALYTICS 154 (2019); *Predictive Policing: Guidance on Where and When to Patrol*, PREDPOL, https://www.predpol.com/how-predictive-policing-works/ (last visited May 7, 2021) ("PredPol uses ONLY 3 data points—crime type, crime location, and crime date/time – to create its predictions. No personally identifiable information is ever used. No demographic, ethnic or socio-economic information is ever used. This eliminates the possibility for privacy or civil rights violations seen with other intelligence-led policing models.").

19. *Are We at the Tipping Point in Police-Community Relations?*, PREDPOL (Jun 11, 2020, 12:02 PM), https://blog.predpol.com/are-we-at-a-tipping-point-in-police-community-relations (purporting that "objective, agreed-upon facts" arising out of data-driven policing can be used to provide transparency in decision making, auditability, and room for discussion around race and policing).

The reality falls far short of the promise. These tools, as designed and deployed in the current legal framework, fail to correct or upend the racial inequity that pervades the criminal legal system. Algorithmic tools aimed at forecasting the behavior of those who are ensnared by the carceral state ensure that all reform efforts will focus on changing the behavior of those being consumed by the system rather than the operation of the system. By choosing to target those who are accused and captured, algorithmic tools presuppose that the people going through the system must be fixed or corrected in some way, rather than altering the system itself. They foster retail reforms where wholesale change is needed. To make matters worse, the prevailing legal regime for rooting out racial bias in criminal legal system decision making insulates these tools from review or intervention, preserving the status quo. At best, they reflect the world around us. At worst, they perpetuate "the New Jim Code," the term given to "new technologies that reflect and reproduce existing inequities" while being "promoted and perceived as more objective or progressive than the discriminatory systems of a prior era."[20]

Reformers who seek to use these tools in the criminal legal system can and should only do so when they design, deploy, and implement them with a basic understanding of the nature of racial inequality. This idea requires that their proponents keep a fundamental truth in mind. That truth, which American history verifies, is that "[r]acial equality is, in fact, not a realistic goal."[21] Simply put, racial inequality is a permanent feature of the institutions that govern us and the society within which we exist. Or to put it in terms that technologists are likely to understand, racism is a feature, not a bug of American life. It is woven into the fabric of our country.

Accordingly, "[e]ven those herculean efforts we hail as successful will produce no more than temporary 'peaks of progress,' short-lived victories that slide into irrelevance as racial patterns adapt in ways that maintain white dominance."[22] To the extent we hope to see more peaks of progress during our lifetimes than valleys of despair, we would do well to accept this premise as true and respond accordingly. That means that rather than attempt to solve or eradicate racism, we should account for the role that racism plays as we design, implement, and engage in oversight of these tools. The evolving policy debate on the use of algorithmic tools provides us with an opportunity to do just that.

20. RUHA BENJAMIN, RACE AFTER TECHNOLOGY: ABOLITIONIST TOOLS FOR THE NEW JIM CODE 10 (2019) (emphasis omitted).

21. Derrick Bell, *Racial Realism*, 24 CONN. L. REV. 363, 363 (1992).

22. *Id.* at 373 (emphasis omitted).

This Article proceeds in three parts. Part I explores the basic nature, character, and history of algorithmic tools across various stages of the criminal legal system, including an accounting of how they are designed, how they work, and the interplay between racial justice and the use of the tools.[23] It complements existing scholarship exposing and addressing the racial justice and fairness concerns the tools raise.[24] It also builds on my own efforts to grapple with the intersection of race and technology[25] by

23. *See infra* Part I.

24. *See* Southerland, *supra* note 11 at 3, 22–25 (detailing gathering of leading experts to foster insights on algorithmic risk assessment and race); Ngozi Okidegbe, *The Democratizing Potential of Algorithms*, 53 CONN. L. REV. (forthcoming 2021); Laura M. Moy, *A Taxonomy of Police Technology's Racial Inequity Problems*, 2021 U. ILL. L. REV. 139 (2019); Rashida Richardson, *Government Data Practices as Necropolitics and Racial Arithmetic*, DATA AND PANDEMIC POLITICS (Oct. 8, 2020), https://globaldatajustice.org/covid-19/necropolitics-racial-arithmetic; *The Bias Embedded in Algorithms*, POCKET (June 18, 2020), https://blog.getpocket.com/2020/06/the-bias-embedded-in-algorithms/ (collecting sources on technology and bias); Nicolás Rivero, *The Influential Project That Sparked the End Of IBM's Facial Recognition Program*, QUARTZ (June 10, 2020), https://qz.com/1866848/why-ibm-abandoned-its-facial-recognition-program/ (detailing how Timnit Gebru, Joy Buolamwini, and Inioluwa Raji have shaped policy through their research); *Some Essential Reading and Research On Race and Technology*, MACHINE (June 2, 2020), https://venturebeat.com/2020/06/02/some-essential-reading-and-research-on-race-and-technology/ (collecting research on race, technology, and bias); Marie Hicks, *Fixing Tech's Built-In Bias*, AMERICAN SCIENTIST, https://www.americanscientist.org/article/fixing-techs-built-in-bias (reviewing literature on race and technology); UCLA Center for Critical Internet Inquiry, *Essential Books by Black Scholars on Technology, Science, and Race*, https://www.c2i2.ucla.edu/racial-justice-and-tech/ (collecting sources); Rachel Courtland, *Bias Detectives: The Researchers Striving to Make Algorithms Fair*, NATURE (June 20, 2018), https://www.nature.com/articles/d41586-018-05469-3 (identifying researchers engaged in algorithmic justice work); Sandra G. Mayson, *Bias in, Bias Out*, 128 YALE L.J. 2218 (2019); Megan Stevenson, *Assessing Risk Assessment in Action*, 103 MINN. L. REV. 303, 341 (2018); Brandon Buskey & Andrea Woods, *Making Sense of Pretrial Risk Assessment*, CHAMPION, June 2018; John Logan Koepke & David G. Robinson, *Danger Ahead: Risk Assessment and the Future of Bail Reform*, 93 WASH. L. REV. 1725 (2018); Sonja Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803 (2014).

25. *See* SOUTHERLAND, *supra* note 11; Vincent Southerland, *With AI and Criminal Justice, the Devil is in the Data*, ACLU BLOG (Apr. 9, 2018, 11:00 AM), https://www.aclu.org/issues/privacy-technology/surveillance-technologies/ai-and-criminal-justice-devil-data. From 2018-2019, I served on an expert Research Advisory Council to provide guidance and consultation as New York City redesigned its pretrial release assessment instrument. Luminosity & the University of Chicago's Crime Lab New York, *Updating the New York City Criminal Justice Agency Release Assessment* 1, 2, 43–44 (June 2020), https://www.nycja.org/assets/Updating-the-NYC-Criminal-Justice-Agency-Release-Assessment-Final-Report-June-2020.pdf. I also served on New York City's Automated Decision Systems Task Force, the first of its kind in the United States, which was established in 2018 and charged with recommending a process for reviewing the City's use of automated decision systems. *See New York City Automated Decision Systems Task Force Report*, N.Y.C. AUTOMATED DECISION SYS. TASK FORCE, (2019), https://www1.nyc.gov/assets/adstaskforce/downloads/pdf/ADS-Report-11192019.pdf; Rashida Richardson, Jason M. Schultz, & Vincent M. Southerland, *Litigating Algorithms 2019 US Report: New Challenges to Government Use of Algorithmic Decision Systems*, AI NOW INST. (Sept.2019). https://ainowinstitute.org/litigatingalgorithms-2019-us.html.

underscoring a truth common to the current menu of algorithmic tools: that if we proceed along our present course, we can at best expect the reification of the pervasive inequities of today.

Part II addresses the potential solutions to the concerns raised by risk assessments in the criminal system.[26] It expands on a growing body of scholarship that grapples with the intersection of race, algorithmic tools, and the law,[27] to produce a series of policy recommendations for how we design, deploy, and assess technological tools in the criminal system. Those policy recommendations include acknowledging the permanence of racism; putting the onus on system actors and tool designers to demonstrate that they do not perpetuate racial harms, regardless of the intent of those who seek to use them; turning the tools on the actors in the system to scrutinize their behavior; and emphasizing qualitative narratives over quantitative data as we press for a system of individualized justice that values the dignity of those facing its punishing power.

Part III concludes with a discussion of the implications of using a racial justice framework and the interventions I have suggested.[28] The recommendations set forth in this Article proceed from the premise that algorithmic tools have the potential to do just as much, if not more, harm than good. Immediate abolition of them or the system in which they operate is unlikely. But the implementation of the recommendations has the potential to bring us one step closer to a criminal legal system radically different than the one we currently employ. Such a system is one in which we have chosen to divest from policing, jails, prisons, and punishment and to invest in education, employment, health, and social welfare. That amounts to a transformation of our criminal system rather than a reform of it.

I. ALGORITHMIC TOOLS IN THE CRIMINAL LEGAL SYSTEM

To begin the examination of the intersection of race and algorithmic tools, it is important to explore the suite of normative concerns and practical challenges that the tools raise at various stages of the criminal system. That focus will unearth the problems presented by the data the tools rely on, the targets that those who traditionally wield them choose, and the critical

26. *See infra* Part II.

27. *See generally* Sean Hill, *Bail Reform & the (False) Racial Promise of Algorithmic Risk Assessment*, UCLA L. REV. (2021) (forthcoming) (applying a racial justice framework rooted in critical race theory to analyze pretrial risk assessments and bail reform in New York and California); BENJAMIN, *supra* note 20 (applying and synthesizing critical race theory and algorithmic tools); Dorothy E. Roberts, *Digitizing the Carceral State*, 132 HARV. L. REV. 1695 (2019) (analyzing the role of race, big data, automation, and computerized prediction in the criminal legal system).

28. *See infra* Part III.

questions that the tools fail to contemplate. By considering these problems in the context of the tools being used in policing, pretrial decision-making, and sentencing, this Part will offer an analytical frame to explore how the theoretical problems play out in practice.[29] Ultimately, this Part demonstrates that if we continue to use these tools in their current configuration, we will only succeed in replicating the bias, racism, and inequity that currently characterizes and consumes the criminal legal system.

A. Brief Introduction to Algorithmic Tools

We begin with a working definition of algorithmic tools. As used in this Article, this term refers to any tools that use statistical data related to past behavior and other relevant traits to predict present or future criminal behavior with the objective of informing decisionmakers about the appropriate criminal legal system outcome or response.[30]

A helpful distinction can be drawn between two sets of tools—predictive tools, which attempt to forecast a particular event or outcome, and surveillance tools, which are used to monitor people, places, and things. The focus of this Article is on predictive tools, which fall within the larger field of predictive analytics: "the use of statistically analyzed data to predict future outcomes."[31] This feature—the analysis and use of group-level data to

---

29. Two points are worth raising. First, while I discuss the problems with the tools at specific stages, those problems are not at all limited to those stages. Each stage provides a lens through which we can see how algorithmic tools operate in practice. It is very much the case that the problems with the design and use of an actuarial tool in, for example, policing may present themselves at bail or sentencing. Second, it is also true that in their deployment and implementation, the tools that I discuss produce additional problems that are worthy of attention. Accordingly, the concerns I have raised are not exhaustive but are intended to capture the broader challenges that the tools present.

30. The term algorithmic tool encompasses what are commonly known as actuarial risk assessments, predictive instruments that use "statistical rather than clinical methods on large datasets of criminal offending rates" and other data deemed relevant to the decision-making process "to determine different levels of offending" or behavior "associated with one or more group traits" BERNARD E. HARCOURT, *Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age* 3 (UNIVERSITY OF CHICAGO PRESS, 2008); *see also* Huq, *supra* note 11, at 1060 ("Algorithmic criminal justice . . . is the application of an automated protocol to a large volume of data to classify new subjects in terms of the probability of expected criminal activity and in relation to the application of state coercion."); Mayson, *supra* note 24, at 2228 (referring to "criminal justice risk assessment" as "the actuarial assessment of the likelihood of some future event, usually arrest for crime."); John Logan Koepke & David G. Robinson, *Danger Ahead: Risk Assessment and the Future of Bail Reform*, 93 WASH. L. REV. 1725, 1752 (2018) ("Typically, risk assessment tools use data about groups of people, like those who have been arrested or convicted, to assess the probability of future behavior.").

31. Jessica M. Eaglin, *Predictive Analytics' Punishment Mismatch*, 14 I/S: J/L & POL'Y FOR INFO SOC'Y 87, 87 (2017).

forecast individual behavior—is common to all predictive algorithmic tools[32] across the criminal legal system.

Algorithmic tools carry with them the promise that they will inform and improve decision-making by the actors employing them. Naturally, discretion provides an entry point for biases to operate, producing unfair outcomes that flow from those biases. Given the centrality of race to the critiques often leveled at the criminal legal system, it should come as no surprise that proponents of algorithmic tools justify their development and use, in part, because they seek to confront and eradicate systemic racial bias and curb biased decision-making. Accordingly, the tools are marketed as race neutral—free from the biases that plague human decision-making,[33] ultimately yielding decisions that are free from bias. Practice, theory, and history paint a different picture—one that is worth confronting if we are ever to advance justice.

What follows is an accounting of the development and use of these tools in policing, pretrial justice, and sentencing.[34] That accounting is framed by the concerns these tools raise: specifically that they yield biased forecasts because they utilize biased data; that they are aimed at those already targeted by the criminal legal system rather than actors in it; and that they encourage profiling.

B. Algorithmic Tools and Policing

1. Theory

The first iteration of algorithmic tools in policing traces back to the twentieth century and the rise of "environmental criminology," which

32. *See supra* notes 29 and 30.

33. The vendors themselves are consistent sources of this marketing. *Predictive Policing: Guidance on Where and When to Patrol*, PREDPOL http://www.predpol.com/how-predictive-policing-works/ (last visited Jan. 8, 2020) ("No demographic, ethnic or socio-economic information is ever used. This eliminates the possibility for privacy or civil rights violations seen with other intelligence-led policing models."); *Predictive Policing Research Breaks New Ground in Philadelphia*, YAHOO! FIN., (Oct. 17, 2013), https://finance.yahoo.com/news/predictive-policing-research-breaks-ground-130000154.html ("Not only does HunchLab enable the combination of many data sources and the intelligent use of temporal patterns, but the new predictive model also enables the prioritization of patrols based on societal impact and local priorities. This data-driven process leads to resource allocations that accurately and fairly reflect societal priorities for public safety, unbiased by neighborhood affluence, race or ethnicity."); *COMPAS Risk & Needs Assessment System*, NORTHPOINTE (2012) http://www.northpointeinc.com/files/downloads/FAQ_Document.pdf (last visited Jan. 8, 2020) ("The tool works well between genders and ethnicities.").

34. *See supra* note 30 (providing explanations of how algorithmic tools work).

focused on the "geography of crime."[35] The idea was to identify and map patterns of criminal behavior to inform policing.[36] Over time, the same maps evolved into digital maps of reported crimes using historical crime data.[37] Police departments eventually hired crime analysts to synthesize crime data and to assist law enforcement with the deployment of limited policing resources.[38]

William Bratton, the Commissioner of the New York City Police Department ("NYPD"), and Jack Maple, the NYPD's Deputy Commissioner for Crime Control Strategies,[39] pioneered data-centered policing.[40] The two developed CompStat, which allowed police leadership to examine reported crime statistics and engage in targeted enforcement to address and reduce crime, measured by arrest rates.[41] These tactics grew out of concerns about systemic corruption in the NYPD and political pressure to address high levels of crime.[42] Data, law enforcement policymakers thought, fostered accountability and professionalism, while reducing crime.[43] When Bratton

---

35. Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U.L. REV. 1109, 1123 (2017).

36. *Id.* at 1123.

37. *Id.* at 1124.

38. *Id.*

39. Bratton served two terms as NYPD Commissioner, from 1994 to 1996 and from 2014 to 2016. Al Baker & J. David Goodman, *Bratton, Who Shaped an Era in Policing, Tries to Navigate a Racial Divide*, N.Y. TIMES (July 25, 2016), https://www.nytimes.com/2016/07/26/nyregion/william-bratton-new-york-city-police-commissioner.html. The development of data-centered policing took place during his first term. *Id.* Maple, for his part, in the 1980s "mapped every train and train station in New York City." He would then use "crayons to mark every violent crime, robbery and grand larceny that occurred," indicating the solved and the unsolved. *Predictive Crime Fighting*, IBM, https://www.ibm.com/ibm/history/ibm100/us/en/icons/crimefighting/ (last visited January 8, 2020). "Using these maps, police officers knew which neighborhoods were being hit by what crimes, and could more efficiently patrol and assist those areas. It was an important first step toward consolidating police data into a tool for crime analysis and prevention." *Id.* The post 9/11 era has seen a proliferation of data-centered policing, as traditional policing expanded to include anti-terrorism and sought to collect and analyze data in keeping with that mission. Paul Hamrick, *Fighting Crime with Data: Law Enforcement in the 21st Century*, FORENSIC FOCUS, (July 5, 2019), https://www.forensicfocus.com/articles/fighting-crime-with-data-law-enforcement-in-the-21st-century/.

40. ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT 29 (2017).

41. *Id.*

42. *Id.*; *see also Predictive Crime Fighting*, *supra* note 39.

43. FERGUSON, *supra* note 40, at 30. It is worth noting that a drop in New York City's crime rate did coincide with the adoption of CompStat, though it is unclear how much CompStat contributed to that decline. Chris Smith, *The Controversial Crime-Fighting Program That Changed Big-City Policing Forever*, N.Y. MAG., (Mar. 2018), https://nymag.com/intelligencer/2018/03/the-crime-fighting-program-that-changed-new-york-forever.html. While "CompStat has helped drive down the city's crime rates to historic lows and revolutionized policing around the world," it also

moved from the NYPD to the Los Angeles Police Department ("LAPD") in 2002, he brought his CompStat approach to another police department reeling from scandal, fraud, and corruption.[44]

Thus began the first efforts to develop predictive policing technologies. Working with academics at area universities, the LAPD experimented with an algorithm to forecast the locations of potential criminal activity.[45] In practice, the analysts fed the algorithm historical crime data to predict the likely location of criminal activity.[46] The program focused on property crimes—specifically burglary, automobile theft, and theft of items from automobiles.[47] A seemingly objective set of considerations informed that focus. First, this suite of crimes generated concern over public safety, tended to be reported and were, therefore, measurable. They could also be addressed by policing practices; they arose from "environmental vulnerabilities" that policing could remedy, and an increased police presence could operate as a deterrent.[48]

The algorithm produced forecasts of criminal activity in geographically precise areas.[49] Police received maps of those areas and instructions to visit them as often as practicable while on patrol.[50] Criminological theory informed practice—resting on the notion that property crimes tend to spread like viruses, either because the environment encourages them or because the same people return to commit them again.[51] Additional variables, like the weather, time of day, proximity to an event, or seasons, provided additional data points for prediction.[52]

Property crime prediction proved to be just the starting point. Two additional versions of predictive policing were developed. The place-based, property-crime-focused iteration of predictive policing evolved to target violent crime.[53] Driving this evolutionary change was the theory that violent crime is the product of particular environmental conditions—a dimly lit alley, proximity to potential victims, gang-related disputes for control over specific

---

fueled a stop and frisk policing regime that led to the harassment of countless New Yorkers of color, driving "considerable debate on just how much credit CompStat, and the NYPD in general, deserves for the crime decline." *Id.*

44. *See* FERGUSON, *supra* note 40, at 29.
45. Ferguson, *supra* note 35, at 1126.
46. *Id.* at 1127.
47. *Id.*
48. *Id.* at 1126–27.
49. *Id.* at 1127.
50. *Id.*
51. *Id.* at 1128
52. *Id.* at 1129.
53. *Id.* at 1132.

territory.[54] Thus, like place-based, property-crime-focused predictive policing, this iteration relied on the notion that "place-based environmental vulnerabilities exist that encourage violent crime, and thus should create a higher risk that crime will occur in that location."[55]

The third iteration of predictive policing represented more of a transformational change. Police began to use "predictive technologies to identify individuals and groups involved in predicted criminal activity."[56] Like the relationships between crime and environmental factors that undergird place-based systems, person-based systems rest on the notion that "negative social networks . . . can encourage criminal activity."[57] This third mode of predictive policing assumes that a small portion of the population possesses an elevated risk of becoming the victim or perpetrator of violence, and that these individuals can be mapped out as a social network to be pinpointed, marked, and surveilled.[58] The result is a shift from "hot spots" where crime might occur to "hot people" who may engage in (or be victims of) violence.[59] Technological advances allowed for intelligence collection and surveillance of suspected individuals and criminal networks, eventually leading to interventions by law enforcement that range from warnings of harsh punishment for targets to increased surveillance.[60]

2. Practice

The record on predictive policing technology is mixed at best. An accounting of initial success in property crime reductions in several California cities—such as Los Angeles, Santa Cruz, Alhambra, and Modesto, along with positive results in Seattle and Atlanta—have been undermined by tests that showed inconclusive results or spikes in crime following initial drops.[61] Boston saw a reduction in violent crime after policing targeted locations where shootings were more likely to take place.[62] Likewise, the city of New Orleans saw a steep decline in its homicide rate after implementing a strategy to target and investigate a cohort of individuals with the highest risk of being involved in gun violence.[63]

54. *Id.* at 1132–33.
55. *Id.* at 1137.
56. *Id.*
57. *Id.*
58. *Id.* at 1138.
59. *Id.* at 1140.
60. *Id.* at 1140–43.
61. *Id.* at 1130.
62. *Id.* at 1134.
63. *Id.* at 1142. The New Orleans predictive policing experiment ended in 2018, when New Orleans Mayor Mitch Landrieu declined to renew the city's partnership with Palantir, a Palo Alto

Yet even these modest successes must be weighed against the potential harm that flows from the use of these tools. In 2016, the Human Rights Data Analysis Group ("HRDAG") reproduced the algorithm utilized by PredPol, a predictive policing software that dozens of police departments nationwide have adopted.[64] PredPol's software consults historical crime data to forecast particular areas—so-called hotspots—that officers should target on a given day.[65] The HRDAG researchers inputted crime data from Oakland, California in order to use the PredPol software to forecast potential drug crime.[66] In response, the algorithm advised the police to target low-income neighborhoods of color, despite concurrent evidence from public health data that drug use is more evenly dispersed throughout the city, and that policing should likewise be more evenly dispersed.[67] This disparity, HRDAG contended, is because officer explicit and implicit biases rooted in race about who to stop, search, and arrest, plagued the records utilized to inform the data, such that the algorithm almost necessarily reproduces accumulated patterns of biased over-policing.[68] Thus, when informed by discriminatory data, the algorithm will work to encourage similarly discriminatory police behavior.[69]

---

based data analytics company. Jonathan Bullington & Emily Lane, *New Orleans ends its relationship With Tech Firm Palantir, Landrieu's Office says*, TIMES-PICAYUNE (Mar. 14, 2018), https://www.nola.com/crime/index.ssf/2018/03/palantir_new_orleans_gang_case.html. Unbeknownst to the New Orleans City Council or the public, Palantir had been operating in New Orleans for six years. Ali Winston, *Palantir Has Secretly Been Using New Orleans to Test Its Predictive Policing Technology*, VERGE (Feb. 27, 2018), https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd.

64. Katia Savchuk, *Justice by The Numbers: Meet the Statistician Trying to Fix Bias in Criminal Justice Algorithms*, PACIFIC STANDARD (Feb. 1, 2019), https://psmag.com/social-justice/justice-by-the-numbers-meet-the-statistician-trying-to-fix-bias-in-criminal-justice-algorithms.

65. *Id.*; *see also Overview*, PredPol, https://www.predpol.com/about/ (PredPol's software "identif[ies] the times and locations where specific crimes are most likely to occur . . . based on on victimization information.").

66. Kristian Lum & William Isaac, *To Predict and Serve?*, 13 SIGNIFICANCE 14, 17 (2016). The results of Lum and Isaac's study have been saddled with the qualification that PredPol is not used to predict drug crimes. Jack Smith IV, *Crime-Prediction Tool PredPol Amplifies Racially Biased Policing, Study Shows*, MIC (Oct. 9, 2016), https://www.mic.com/articles/156286/crime-prediction-tool-pred-pol-only-amplifies-racially-biased-policing-study-shows. This qualification, while valid, does not detract from what the study demonstrated: "the potential for predictive policing software to perpetuate historical biases in enforcement." William Isaac & Kristian Lum, *Setting the Record Straight on Predictive Policing and Race*, MEDIUM (Jan. 3, 2018), https://medium.com/in-justice-today/setting-the-record-straight-on-predictive-policing-and-race-fe588b457ca2.

67. Lum & Isaac, *supra* note 66, at 17.

68. *Id.* at 15.

69. *Id.*

Person-based predictive systems suffer from similar shortcomings. A RAND Corporation[70] study of the Chicago Police Department's ("CPD") Strategic Subject List ("SSL") is a helpful example. The SSL is "a computerized assessment tool that incorporates numerous sources of information to analyze crime as well as identifies and ranks individuals at risk of becoming a victim or possible offender in a shooting or homicide."[71] Developed by the Illinois Institute of Technology, and utilized by the CPD as early as 2012, this tool assigns risk tiers to individuals based on variables, like an individual's age during their latest arrest, the number of times they have been apprehended for use of an unlawful weapon, and the number of times they have been a victim of aggravated assault and battery.[72] Because the majority of these variables rely upon arrest records rather than actual convictions, however, the SSL runs a high risk of including individuals who have not even committed a crime, and of reflecting the CPD's biased policing practices.[73] Indeed, research demonstrated that the SSL led to increased contact with those who were already in frequent contact with law enforcement.[74] What is worse, the SSL did not reduce gun violence, even as the number of individuals on the list tripled over three years.[75]

---

70. According to its website, "The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous." *About the RAND Corporation*, RAND CORP., https://www.rand.org/about.html (last visited June 1, 2021). RAND describes its history as follows: "On May 14, 1948, Project RAND—an organization formed immediately after World War II to connect military planning with research and development decisions—separated from the Douglas Aircraft Company of Santa Monica, California, and became an independent, nonprofit organization. Adopting its name from a contraction of the term *research and development*, the newly formed entity was dedicated to furthering and promoting scientific, educational, and charitable purposes for the public welfare and security of the United States." *History and Mission*, RAND CORP., https://www.rand.org/about/history.html (last visted June 1, 2021).

71. Rashida Richardson et al., *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, 94 N.Y.U. L. REV. ONLINE 15, 31 (2019)

72. *Id.*

73. An arrest of a particular nature and character does not always yield a conviction of the same nature and character for the individual arrested, especially in those instances when law enforcement authorities engage in biased policing. *Id*. at 28–29, 29 n.57 (citing a Department of Justice investigatory report that found the Chicago Police Department's pattern or practice of unconstitutional conduct resulted in false arrests and convictions of incalculable proportion).

74. David Robinson & Logan Koepke, UPTURN, STUCK IN A PATTERN: EARLY EVIDENCE ON "PREDICTIVE POLICING" AND CIVIL RIGHTS 9 (2016), https://www.upturn.org/static/reports/2016/stuck-in-a-pattern/files/Upturn_-_Stuck_In_a_Pattern_v.1.01.pdf.

75. *Id.* Notably, the Chicago Police Department decommissioned the SSL in January 2020 following a report by the Office of Inspector General detailing myriad problems with the program. Sam Charles, *CPD Decommissions 'Strategic Subject List'*, CHI. SUN –TIMES (Jan. 27, 2020, 2:11pm), https://chicago.suntimes.com/city-hall/2020/1/27/21084030/chicago-police-strategic-subject-list-party-to-violence-inspector-general-joe-ferguson. Those problems included "the unreliability of risk scores and tiers; improperly trained sworn personnel; a lack of controls for internal and external access; interventions influenced by . . . risk models which may have attached

Significant harms can flow from an algorithmic tool that targets policing in particular communities and suggests repeatedly returning to those communities.[76] For example, more interactions between Black people and police make Black people vulnerable to violence at the hands of law enforcement; increases the likelihood of arrest; and fosters likely involvement with the criminal legal system, driving up rates of arrest and incarceration.[77] Repeated exposure to police tends to increase the vulnerability of those policed to "violence-producing insecurities" that officers experience during encounters.[78] Finally, Black people who come into frequent (and unwarranted) contact with law enforcement develop a decreased perception of police legitimacy, which can cause them to "resist police authority, assert rights, or flee upon seeing or encountering the police, each of which increases the likelihood of police violence."[79]

Despite the real world harms these tools can produce, it is hard to argue with the use of technology when success is defined as less crime, more cases cleared, and a greater sense of public safety for some segment of society. On those terms, even the minimal success of these tools allows justice actors who seek to use them to ignore a number of questionable assumptions under the veneer of a technological solution.[80] Chief among those assumptions is one of the basic vulnerabilities of all actuarial risk assessments raised by the problems revealed through studies of predictive policing tools: bad data.

---

negative consequences to arrests that did not result in convictions; and a lack of a long-term plan to sustain the . . . models." *Id.* (quoting Office of Inspector General) (internal quotation marks omitted).

76. *See* Devon W. Carbado, *Blue-on-Black Violence: A Provisional Model of Some of the Causes*, 104 GEO. L.J. 1479, 1509 (2016) (noting that heightened police interactions with Black communities not only reflects, but also reinforces racial stereotypes of Black people as violent and dangerous).

77. *Id.* at 1508–11.

78. *Id.* One example of this phenomenon is "'masculinity threat,' which is an officer's sense that his masculinity is being undermined or challenged during an interaction." *Id*. Officers who experience this phenomenon are, on balance, more likely to deploy violence than those who do not. *Id*; *see also* L. Song Richardson & Phillip Atiba Goff, *Interrogating Racial Violence*, 12 OHIO ST. J. CRIM. L. 115, 128–42 (2014) (defining and discussing masculinity threat and its relationship to racial violence).

79. Carbado, *supra* note 76, at 1511; *see also* CIVIL RTS. DIV., U.S. DEP'T OF JUSTICE, INVESTIGATION OF THE BALTIMORE CITY POLICE DEPARTMENT (2016), at 7.

80. In June 2020, the city of Santa Cruz, California, recognizing the concerns raised by predictive policing, became the first American city to bar its use. Nicholas Ibarra, *Santa Cruz, Calif., Bans Predictive Policing Technology*, GOV'T. TECH. (June 24, 2020), https://www.govtech.com/public-safety/Santa-Cruz-Calif-Bans-Predictive-Policing-Technology.html.

3. Critique: Flawed Data as Destiny and Garbage In, Garbage Out

Data is the lifeblood of all predictive technology. In the context of the criminal legal system, data is rife with imperfections and is irreversibly tainted by racism and the social hierarchies it produces and supports.[81] Those indelible flaws are, in large part, the byproduct of the nature of crime data—police do not just use data—they create the data that algorithmic tools and technologies depend on.[82] Thus, police decision-making plays an outsized role in shaping our perceptions of crime and criminal behavior.[83] The vulnerabilities in the data start with simple, innocent, human error: People can make mistakes in data collection, input, integration of datasets, and cleansing to remedy duplicative entries.[84] Data can also be incomplete, as its creation is often wholly dependent on actors within the criminal legal system—both the consumers and the consumed.[85] Everything from the underreporting of crime by communities that have lost faith in law enforcement, or have some other reason not to report crime,[86] to the manipulation of crime statistics[87] by police can produce data that paints an incomplete portrait of a community—and therefore an incomplete and flawed field of vision for a predictive policing tool.

Another source of this flawed data problem, independent from the motivations of the stat-juking officer, emerges from the nature of interactions between police and citizens. Arrest statistics, which mark the point of contact between law enforcement and alleged perpetrators, are not updated to reflect

81. *See* Andrew Guthrie Ferguson, *Illuminating Black Data Policing*, 15 OHIO ST. J. CRIM. L. 503, 504 (2018) (explaining that all big data policing technologies suffer from a lack of transparency, racial bias, and legal uncertainty).

82. Elizabeth E. Joh, *Feeding the Machine: Policing, Crime Data & Algorithms*, 26 WM. & MARY BILL OF RTS. J. 287, 289 (2017) (emphasis omitted).

83. *Id.* at 290.

84. Ferguson, *supra* note 35, at 1145–46.

85. *Id.* at 1146–47; *see also* BARRY FRIEDMAN, UNWARRANTED: POLICING WITHOUT PERMISSION 266–68 (2017) (detailing the potential ways data can be erroneous given how it is gathered).

86. *See* P. Jeffrey Brantingham, *The Logic of Data Bias and its Impact on Place-Based Predictive Policing*, 15 OHIO ST. J. CRIM. L. 473, 475 (2018) (explaining that crime is substantially underreported across crime types by all racial groups, though at varying degrees); *see also* Ferguson, *supra* note 81, at 514–16 (describing differences in crime reporting by communities of color and for particular types of crime).

87. Matt Hamilton, *LAPD Captain Accuses Department of Twisting Crime Statistics to Make City Seem Safer*, L.A. TIMES (Nov. 6, 2017), http://www.latimes.com/local/lanow/la-me-ln-lapd-crime-stats-claim-20171103-story.html; *see also* Brantingham, *supra* note 86, at 475 ("A related source of bias is police intentionally undercounting crime either through intentional mislabeling or failing to report" stemming from "perverse incentives for police to make the world seem better than it actually is.").

how the arrest was resolved by the criminal legal system.[88] Cases that the government dismisses, or those that resolve with a plea on charges less serious than those for which an arrest was made, or those where an accused person accepts a plea to charged conduct that they did not in fact commit, will naturally skew the data and likewise present a distorted picture of when and where crime is occurring and who is responsible for it.[89] For example, an individual may be arrested for a robbery and charged accordingly (or institutional pressures may lead a prosecutor to charge the most serious offense consistent with the facts presented). Ultimately, that case may be resolved with a guilty plea to a lesser charge—such as assault or theft—that more closely aligns with the behavior of the accused. Traditional crime data would reflect the robbery, rather than the ultimate, less serious outcome. What is reflected and read in the data is a community that appears to be dramatically more dangerous than it actually is.

Compounding the concerns raised by these serious shortcomings is the fact that the most pressing data-related problems occur at the intersection of race: biased data.[90] Simply put, "[p]olice data remains colored by explicit and implicit bias. Police data is racially coded, shaded by millions of distrustful looks and thousands of discomfiting physical encounters."[91] A cursory examination of policing practices reveals the pervasive influence of bias—and racial bias in particular—on law enforcement.[92]

---

88. Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 GA. L. REV. 109, 133–34 (2018).

89. *See, e.g.*, Jed S. Rakoff, *Why Innocent People Plead Guilty*, N.Y. REV. (Nov. 20 2014), https://www.nybooks.com/articles/2014/11/20/why-innocent-people-plead-guilty/; Walter Pavlo, *Are Innocent People Pleading Guilty? A New Report Says Yes*, FORBES (July 31 2018, 8:06 PM), https://www.forbes.com/sites/walterpavlo/2018/07/31/are-innocent-people-pleading-guilty-a-new-report-says-yes/?sh=2858edea5193; Jed S. Rakoff, WHY THE INNOCENT PLEAD GUILTY AND THE GUILTY GO FREE (2021).

90. Ferguson, *supra* note 35, at 1148–49.

91. FERGUSON, *supra* note 40, at 131–32.

92. Balko, *supra* note 3 (collecting seventeen studies produced examining data from 2002 detailing racial bias and discriminatory policing). Rooting out misconduct and bias is incredibly challenging, given that there are more than 18,000 law enforcement agencies nationwide. CIVIL RTS. DIV., UNITED STATES DEP'T OF JUSTICE, THE CIVIL RIGHTS DIVISION'S PATTERN AND PRACTICE POLICE REFORM WORK: 1994-PRESENT 1 (Jan. 2017). Since 1994, the Department of Justice's Civil Rights Division has had the authority to investigate and litigate cases involving patterns or practices by law enforcement that violate the Constitution or federal civil rights statutes. *Id.* at 3. Since the Division began that work, it has opened sixty-nine formal investigations and entered into forty reform agreements addressing their investigatory findings. *Id.* *See* Richardson et al., *supra* note 71, at 199–202 (describing how criminal legal system data is reflective of biased police practices).

Behind the deaths of George Floyd, Michael Brown, Eric Garner, Tamir Rice, Philando Castille, Stephon Clark, Pamela Turner, Korryn Gaines,[93] and countless other people of color killed by the police are staggering data points that underscore the racism that pervades policing.[94] Black people are more likely than their white counterparts to be stopped, searched, arrested, and victimized by the police.[95] A 2019 analysis of 100 million municipal and state patrol traffic stops from dozens of jurisdictions nationwide over a decade revealed that Black drivers are 20% more likely to be pulled over than their white counterparts.[96] The same analysis determined that the threshold for searching Black and Latino drivers was lower than that applied to their white counterparts, meaning that searches of Black and Latino drivers were premised on fewer contextual factors that give rise to suspicion than searches of white drivers.[97] For young men of color, police force is among the leading causes of death.[98] About 1 in 1,000 Black men and boys can expect to lose their lives to police violence—a risk 2.5 times higher than that of their white peers.[99] On the whole, Black people are three times more likely to be killed by police.[100] These numbers, along with the incidents they represent, led to investigations by the Department of Justice's Civil Rights Division, which found widespread racially discriminatory policing practices in places like

93. In 2019 alone, 999 people were shot and killed by the police, 249 of whom were Black, 163 of whom were Hispanic (for a total of 367 non-white victims) and 405 of whom were white, with the remainder reported as being of unknown or other races. *Fatal Force*, WASH. POST, https://www.washingtonpost.com/graphics/2019/national/police-shootings-2019/ (Aug. 10, 2020). For a sampling of media reports regarding this phenomena, see *110 Black Men And Boys Killed By Police*, NEWONE (May 5, 2021), https://newsone.com/playlist/black-men-boy-who-were-killed-by-police/item/53; *#SayHerName: Black Women And Girls Killed By Police*, NEWSONE (Oct. 14, 2019), https://newsone.com/playlist/black-women-girls-police-killed-photos/item/1.

94. *See Mapping Police Violence*, https://mappingpoliceviolence.org/ (last updated February 16, 2021) (providing a comprehensive statistical examination of police violence and deaths at the hands of law enforcement since 2013).

95. Elizabeth Hinton et al., *An Unjust Burden: The Disparate Treatment of Black Americans in the Criminal Justice System*, VERA INST. JUST. (May 2018), https://storage.googleapis.com/vera-web-assets/downloads/Publications/for-the-record-unjust-burden/legacy_downloads/for-the-record-unjust-burden-racial-disparities.pdf.

96. Emma Pierson et al., *A Large-Scale Analysis of Racial Disparities in Police Stops Across The United States*, 4 NATURE HUMAN BEHAVIOR 736, 737 (2019), https://www.nature.com/articles/s41562-020-0858-1.

97. *Id.* at 6.

98. Frank Edwards et. al., *Risk of Being Killed by Police Use of Force in the United States by Age, Race–Ethnicity, and Sex*, 116 PROC. NAT'L ACAD. SCI. 16,793, 16,793 https://www.pnas.org/content/116/34/16793.

99. Amina Khan, *Getting Killed by Police is a Leading Cause of Death for Young Black Men in America*, L.A. TIMES (Aug. 16, 2019), https://www.latimes.com/science/story/2019-08-15/police-shootings-are-a-leading-cause-of-death-for-black-men.

100. Mapping Police Violence, *supra* note 94.

Ferguson, Missouri; Newark, New Jersey; Baltimore, Maryland; New York, New York; and Chicago, Illinois.[101]

Given the racialized nature of policing, it should come as no surprise that law enforcement practices have generated biased data. Reliance on biased data by predictive policing tools has the potential to produce devastating consequences.[102] Predictive policing tools "look[ ] at crime in one geographic area, incorporate[ ] it into historical patterns,"[103] and deliver a prediction that often justifies a continued or increased police presence in a particular community.[104] The effect is twofold.[105] First, targeting of law enforcement resources in a specific community based on past policing patterns may lead to more arrests of individuals in that community, giving the impression that members of that community are more likely to engage in criminal behavior.[106] Second, the mere presence of law enforcement guarantees an increase in arrests, and, in turn, the creation of more bad data. The result is a "pernicious feedback loop", where "[t]he policing itself spawns new data, which then justifies more policing."[107]

In other words, human fallibilities that track racial inequities taint the precise data on which we focus these tools. For example, a host of factors feed into the discretion exercised by officers deciding whether to make a stop and arrest. Those factors might relate to the dynamics of the interaction between suspect and officer. The wishes of a complainant can affect both the decision to charge and the nature of the charge. The incentives for increased or decreased enforcement affect officers' decisions about formal intervention versus informal resolution of misconduct.[108] These variables shape the data

101. For a summary of the findings of DOJ's work in these jurisdictions, *see* CIVIL RTS. DIV., *supra* note 92.'

102. Indeed, civil rights leader Roy Wilkins, the Executive Secretary and Executive Director of the National Association for the Advancement of Colored People from 1955 through 1977, expressed this precise concern with the advent of computers in a 1967. *See* MCILWAIN, *supra* note 6, at 242 ("He knew that white America associated black people with crime. He was afraid that that association, and data that confirmed it, would be fed into, ingested in, and processed by a powerful new computer system—one that stored, connected, and distributed large amounts of decision-driving data that could negatively impact black people's lives.").

103. CATHY O'NEILL, WEAPONS OF MATH DESTRUCTION 75 (2016).

104. Ferguson, *supra* note 35, at 1148–49.

105. *See* Brantingham, *supra* note 86 at 475 (describing how implicit bias can affect a place based predictive policing models).

106. Selbst, *supra* note 88, at 134–35.

107. O'Neill, *supra* note 103, at 87.

108. *See* Ekow N. Yankah, *Pretext and Justification: Republicanism, Policing, and Race*, 40 CARDOZO L. REV. 1543, 1580-81 (2019) (noting that the enforcement of traffic violations runs the risk of police officers exercising broad discretion to stop drivers for impermissible reasons such as race).

generated.[109] A mountain of evidence demonstrates that race is one of those variables.[110] It has the capacity to shape everything about police practices, from interactions between officers and citizens to law enforcement priorities.

Even if race is not the principle motivating factor, its influence is reflected in law enforcement data. Machine learning algorithms, which learn how to reproduce the data they are fed, will naturally reproduce that biased data.[111] Predictive systems, then, will identify people and locations that reflect prior police interactions.[112] Thus, despite the fact that "[n]one of the algorithms use race in their model (and in fact strip it out) . . . the technologies end up targeting communities of color."[113] In short, at best, predictive policing tools premised on biased data will reflect that biased data, reinforcing the discriminatory forces and race-based assumptions that produced it in the first place.

To be clear, vendors of predictive policing tools, confronted with the challenges of bad data, have made efforts to cleanse their products of the taint of racism. In some instances, they have done so by relying on data points that do not explicitly rely on race but correlate with it, like zip code or economic status of a particular location.[114] These efforts make the link between racially tainted data and racially tainted forecasts feel, at first glance, like more of a significant risk than a hard and fast reality.[115] The assumption is that if a vendor does not use a data point that is traditionally tied to race, the forecasts produced by the technology will be non-racialized. This is especially true of place-based predictive policing systems. One such vendor, PredPol:

> uses only 3 data points—crime type, crime location, and crime date/time—to create its predictions. No personally identifiable information is ever used. No demographic, ethnic or socio-economic information is ever used. This eliminates the possibility for privacy or civil rights violations seen with other intelligence-led or predictive policing models.[116]

109. Joh, *supra* note 82, at 297–301.

110. Balko, *supra* note 3 (detailing the influence of race and racism on the criminal legal system).

111. *Id.* at 300–01; *see also supra* notes 64–69 and accompanying text; Richardson et al., *supra* note 71,at 192.

112. Joh, *supra* note 82, at 301.

113. Ferguson, *supra* note 81, at 516.

114. FERGUSON, *supra* note 40, at 75.

115. Andrew Guthrie Ferguson, *The Truth About Predictive Policing and Race*, APPEAL (Dec. 7, 2017), https://theappeal.org/the-truth-about-predictive-policing-and-race-b87cf7c070b1/.

116. *The Three Pillars of Predictive Policing*, PREDPOL, https://www.predpol.com/law-enforcement/#predPolicing (last visited Jan. 10, 2020) (emphasis omitted).

Hunchlab generates its forecasts from public reports of crime, supplemented with data about the geography, weather patterns, and things like the locations of community resources.[117]

Unfortunately, these efforts do not fully mitigate the risks of flawed data. Patterns of reported crimes, like policing patterns and nearly everything about the criminal legal system, vary by race.[118] Tools that look to community resources, like the locations of schools, restaurants, liquor establishments, and transportation hubs[119] have to contend with historical, racialized patterns of residential segregation that have produced an uneven geographical distribution of such establishments.[120] For example, if a correlation is drawn between criminal activity and community center locations, and those centers are largely found in public housing residences

---

117. The CEO of the company which developed and sold Hunchlab in January 2019 detailed their efforts to avoid running afoul of civil rights concerns:

> "Forecast places, not people: We would forecast locations with the highest likelihood of a crime at a given point in time. We do not attempt to make predictions about the actions of people. Limit input data to places, not people: We would not use data about people – no arrests, no social media, no gang status, no criminal background information. Reported events: We would generate forecasts based on public reports of crime, not arrests or other data originating in law enforcement activities. Supplement reported data: One way to reduce bias is to draw on multiple sources of data. We knew that we could generate forecasts using just the crime reports, but we believed that by supplementing reported crimes with other relevant data, ideally from independent, open sources, we could mitigate bias in the reporting data. Typical examples might include lighting, school schedules, locations of community infrastructure, weather, or locations of bars."

Robert Cheetham, *Why We Sold Hunchlab*, AZAVEA (Jan. 23, 2019), https://www.azavea.com/blog/2019/01/23/why-we-sold-hunchlab/.

118. Caroline Haskins, *Academics Confirm Major Predictive Policing Algorithm is Fundamentally Flawed*, VICE (Feb. 14, 2019), https://www.vice.com/en_us/article/xwbag4/academics-confirm-major-predictive-policing-algorithm-is-fundamentally-flawed; *see also* Balko *supra* note 4 (reviewing studies that detail racial bias in the criminal legal system).

119. *HunchLab: Under the Hood*, AZAVEA 19–20 (2015), https://cdn.azavea.com/pdfs/hunchlab/HunchLab-Under-the-Hood.pdf; *see also* Maurice Chammah, *Policing the Future*, MARSHALL PROJECT (Feb. 3, 2016), https://www.themarshallproject.org/2016/02/03/policing-the-future#.9vrCo3ZOH.

120. Deborah N. Archer, *The New Housing Segregation: The Jim Crow Effects of Crime-Free Housing Ordinances*, 118 MICH. L. REV. 173, 185 (2019) ("Through exclusionary housing policies that masquerade as race-neutral principles of rational planning and home rule, homogeneous municipalities can, and do, act on their worst biases. Many local communities exercise their local power to relegate poor people of color to marginalized, resource-starved neighborhoods, away from the economic prosperity of their own communities."); RICHARD ROTHSTEIN, THE COLOR OF LAW, xvi (2017); *see also* Danyelle Solomon et. al., *Systematic Inequality: Displacement, Exclusion, and Segregation*, CTR. FOR AM. PROGRESS 4, 10 (Aug. 2019), https://www.americanprogress.org/issues/race/reports/2019/08/07/472617/systemic-inequality-displacement-exclusion-segregation/ ("Racial segregation has contributed to persistent disparities in access to public goods—such as parks, hospitals, streetlights, and well-maintained roads—and has undermined wealth building in communities of color nationwide.").

inhabited by communities of color, the tools will forecast crime to take place in those locations.

At bottom, this is a case of garbage-in, garbage-out. Or as some call it, "racism in, racism out."[121] The solutions most often posited to address flawed data fall short. That is because no solution can fully erase the vulnerabilities of racism, biases, and errors that are embedded in the information used by these instruments to produce their forecasts.[122] As we will see in the following Section, which examines actuarial risk assessment tools and pretrial justice, the problem of flawed data is just the first of several overarching problems with these tools.

C. Algorithmic Tools and Pretrial Justice

1. Theory

For well over a half century, reformers have engaged in efforts to rethink America's pretrial justice system.[123] In its modern form, pretrial justice is best understood as the point in the system following arrest and coinciding with a prosecutor's charging decision. It is at that point when a judge must make a decision about whether to detain an individual, release them from law enforcement custody, or condition a person's release from custody on meeting an obligation, such as paying a monetary amount to ensure a return to court. The origins of America's pretrial system trace back over two centuries ago to English common law, which presumed release for people accused of noncapital crimes barring a serious risk of flight.[124] Over the last half century, the right to bail has evolved in the United States, incorporating an additional consideration of the likelihood that the accused will pose a risk to public safety.[125]

---

121. Stephanie Buranyi, *Rise of the Racist Robots—How AI is Learning All of Our Worst Impulses*, GUARDIAN (Aug. 8, 2017), https://www.theguardian.com/inequality/2017/aug/08/rise-of-the-racist-robots-how-ai-is-learning-all-our-worst-impulses (quoting Hamid Khan, an organizer with the Stop LAPD Spying Coalition).

122. One solution proposed by researchers is "that every dataset be accompanied with a datasheet that documents its motivation, composition, collection process, recommended uses . . ." Timnit Gebru et al., *Datasheets for Datasets*, CORNELL UNIVERSITY, 1 (last revised Mar. 20, 2020) (working paper), https://arxiv.org/abs/1803.09010v7.pdf. This solution has "the potential to increase transparency and accountability . . . mitigate unwanted biases . . . and help researchers and practitioners select more appropriate datasets for their chosen tasks." *Id*. at 2. Nevertheless, it "do[es] not provide a complete solution to mitigating unwanted biases or potential risks or harms." *Id*. at 10.

123. *See* Sandra G. Mayson, *Dangerous Defendants*, 127 YALE L.J. 490, 502–09 (2018) (describing the development of pretrial justice since the 1960s).

124. Crystal Yang, *Toward an Optimal Bail System*, 92 N.Y.U. L. REV. 1399, 1410–11 (2017).

125. *Id*. at 1412.

Three waves of reform have driven the evolution of pretrial justice.[126] The first wave of reform, which provided the foundation for the current pretrial justice regime, culminated in the Bail Reform Act of 1966, signed into law by President Lyndon Johnson.[127] The law was enacted largely in response to a growing chorus of voices decrying the inequities in the system's operation. Judges tended to exercise discretion by setting unaffordable money bail amounts that inevitably relegated the poor to pretrial detention.[128] As then-Attorney General Robert F. Kennedy testified before a Congressional committee:

> [T]he rich man and the poor man do not receive equal justice in our courts. And in no area is this more evident than in the matter of bail. . . . [B]ail has become a vehicle for systematic injustice. Every year in this country, thousands of persons are kept in jail for weeks and even months following arrest. They are not yet proven guilty. They may be no more likely to flee than you or I. But, nonetheless, most of them must stay in jail because, to be blunt, they cannot afford to pay for their freedom.[129]

The Bail Reform Act emphasized "the long-standing objective that bail should be used solely to prevent flight risk," imposing a presumption of release unless doing so would undermine the chance that the accused would not return to court.[130]

The presumption of release and focus on risk of flight shaped bail decisions until the early-to-mid 1980s when concerns about public safety and pretrial crime prompted a dramatic change and a second wave of reform.[131] States passed laws that allowed for preventive detention—the pretrial incarceration of those deemed too dangerous to society to be released.[132] Despite efforts to upend preventive detention, which is rooted in the idea of detaining individuals based on the possibility that they pose a danger to public safety because they may commit some future offense while their criminal case is pending, the Supreme Court upheld the more restrictive

126. *See* Mayson, *supra* note 123, at 502–09 (describing waves of bail reform).

127. Bail Reform Act of 1966, 18 U.S.C. § 3142, *et seq*.

128. Yang, *supra* note 124, at 1412–13.

129. Federal *Bail Procedures*: *Hearings Before the Subcomm. on Constitutional Rights and Subcomm. on Improvements in Judicial Mach. of the Comm. on the Judiciary* 88th Cong. 27(1964) (statement of Hon. Robert F. Kennedy, Att'y Gen. of the United States) [hereinafter *Bail Legislation*].

130. Yang, *supra* note 124, at 1413.

131. *See* Lauryn P. Goldin, *Disentangling Flight Risk from Dangerousness*, 2016 B.Y.U. L. Rev. 837, 847–52 (2016) (describing shift in statutory language to consider dangerousness in bail determinations); *United States v. Salerno*, 481 U.S. 739, 742 (1987) (describing consideration of dangerousness in federal bail statute as reaction to pretrial crime).

132. Yang, *supra* note 124, at 1413.

pretrial regime in *United States v. Salerno.*[133] *Salerno* marked a sea change in pretrial justice, as states nationwide enacted bail legislation that allowed courts to explicitly consider the danger the accused poses to the public.[134]

Bail reform has come full circle. The fear that drove the first wave of bail reform—that wealth determined who would be freed pretrial—has animated the latest series of reform efforts. Jurisdictions nationwide have been prodded by litigation and advocacy to replace their cash-bail-based pretrial systems with risk-based systems that employ algorithmic tools called pretrial risk assessments to guide release and detention decisions.[135] The adoption and development of pretrial risk assessments was sparked half a century ago by the Manhattan Bail Project, which consisted of a collaboration between New York City's criminal courts and the Vera Institute of Justice.[136] The Manhattan Bail Project introduced the use of a formal questionnaire in the pretrial process to elicit information about an accused's personal characteristics and family and community ties that could be assigned point values in order to determine whom the courts could safely release pretrial without bail.[137] The data produced by the Vera effort:

> [P]rovided objective factors to be used in setting release conditions. Scoring each community link and requiring a threshold score for release on one's 'own recognizance' created a crude but functional actuarial instrument for risk assessment, replacing the essentially clinical judgment of a judge who set financial terms on the basis of a holistic but subjective evaluation.[138]

Today, approximately forty jurisdictions in twenty-eight states use some form of pretrial risk assessment instrument.[139] Each of these tools aims to

133. 481 U.S. 739 (1987).

134. Yang, *supra* note 124, at 1414–15.

135. Mayson, *supra* note 123, at 508–09.

136. *Manhattan Bail Project*, VERA INST. OF JUSTICE, https://www.vera.org/publications/manhattan-bail-project-official-court-transcripts-october-1961-june-1962 (last visited Mar. 13, 2021).

137. *Preventive Detention in New York: From Mainstream to Margin and Back*, CTR. ON THE ADMIN. OF CRIM. L. 4–6 (2017), https://www.law.nyu.edu/sites/default/files/upload_documents/2017-CACL-New-York-State-Bail-Reform-Paper.pdf. Attorney General Kennedy's testimony previewed the current embrace of risk assessments, as his testimony before Congress pointed to the regime put in place by the Manhattan Bail Project as an example of successful bail reform efforts. *Bail Legislation*, *supra* note 129, at 4.

138. Jonathan Simon, *Reversal of Fortune: The Resurgence of Individual Risk Assessment in Criminal Justice*, 1 ANN. REV. L. & SOC. SCI. 397, 406 (2005).

139. Mayson, *supra* note 123, at 510; PRETRIAL JUST. INST., SCAN OF PRETRIAL PRACTICES 2019 25 (2019) [hereinafter SCAN OF PRETRIAL PRACTICES], https://university.pretrial.org/viewdocument/scan-of-pretrial-practices-pji-20 ("[I]n 2017, approximately one in four people in the United States lived in a jurisdiction that employed a validated evidence-based pretrial assessment tool, up from one in 10 people in 2013.").

predict who, among the accused, is at risk of being rearrested or failing to appear in court. Some jurisdictions developed these tools on their own, while private corporations, foundations, academics, and data scientists developed and designed others independently for adoption and use by a jurisdiction.[140] The factors assessed by the tools vary, but prior convictions and pending charges are commonly utilized.[141] A checklist tool—one that a pretrial services agency or court authority administers and determines the presence of a list of factors or characteristics—is the most widely used methodology.[142] Statisticians analyze aggregated pretrial data to determine the characteristics or traits of an accused person that most closely correlate with the outcome to be assessed by the tool.[143] Tool makers assign points to those characteristics or traits—called risk factors—that correspond to the relationship between the factor and the outcome. A risk score is calculated by determining which risk factors apply to the individual being assessed and adding up that score.[144] Some tools do not reveal the weights—or scores—assigned to individual factors or reveal what factors are being taken into account.[145] Though pretrial tools weigh an individual's risk of re-arrest and failure to appear, "[m]ost of the existing instruments produce a single score that represents the risk of either one occurring."[146]

This merger of risks is problematic for a number of practical and policy reasons. First, dangerousness and flight are distinct concerns that can lead to pretrial detention. Accordingly, the Federal Bail Reform Act and the majority of state bail statutes require that each phenomenon be considered separately.[147] In many states, while detention may be justified by the flight

140. An example of this bespoke design process is the effort undertaken by New York City's Criminal Justice Agency, which is responsible for managing the city's pretrial justice system. *Release Assessment*, N.Y.C. CRIM. JUST. AGENCY, https://www.nycja.org/release-assessment (last visited Jan. 10, 2020).

141. Yang, *supra* note 124, at 1484.

142. *See, e.g.*, *supra* note 139.

143. *See* Koepke & Robinson, *supra* note 30, at 1752–54 (describing how pretrial risk assessments function); DESMARAIS & LOWDER, *supra* note 8 (describing different forms of pretrial risk assessment and its basic mechanics).

144. Mayson, *supra* note 123, at 509.

145. *E.g.*, State v. Loomis, 881 N.W.2d 749, 763–64 (Wis. 2016) (requiring COMPAS to inform courts when the company invokes the proprietary nature of its software to "prevent disclosure of information relating to how factors are weighed or how risk scores are to be determined"). Further, even among pre-trial risk assessment tools that do disclose their weights, many have not been "validated" to show that the algorithm measures what it is intended to measure. Brandon Buskey & Andrea Wood, *Making Sense of Pre-trial Risk Assessments*, CHAMPION 1, 18 (June 2018), https://www.nacdl.org/Article/June2018-MakingSenseofPretrialRiskAsses. Validation studies often do not reveal how data points are weighted or what scores serve as cutoffs for different risk levels. *Id.*

146. Mayson, *supra* note 123, at 509–10.

147. Goldin, *supra* note 131, at 872–84.

risk one presents, a statutorily mandated separate finding is required to actually impose detention.[148] Different conditions of release—electronic monitoring or a stay away order instead of cash bail—may flow from a separate consideration of flight and dangerousness.[149]

From a policy standpoint, merging the two types of risk can lead to an inadvertent overestimation of both. For example, mixing the two may mean that a judge's estimation of flight risk is tainted by fears of one's risk of dangerousness, while estimation of the risk of danger that one may pose may be tainted by fears that someone poses a flight risk.[150] Combining the two forms of risk also prevents judges from understanding and accounting for the importance of each risk on its own to their bail determinations.[151] Notwithstanding the concerns that flow from combining risks, states nationwide have adopted risk assessment instruments to inform pretrial decision-making.[152]

2. Practice

As jurisdictions nationwide adopt pretrial algorithmic tools—commonly known as pretrial risk assessment instruments—as part of their reforms, the efficacy of the tools remains in question.[153] At worst, they carry the potential to reproduce disparity.[154] At best, their introduction is accompanied by decarceratory results without changing the racial

---

148. *Id.* at 873.

149. *Id.* at 881–85, 893–97.

150. *Id.* at 886–88.

151. *Id.* at 892–93.

152. *See* SCAN OF PRETRIAL PRACTICES, *supra* note 139, at 25; Matt Henry, *Risk Assessment: Explained*, APPEAL (Mar. 25, 2019), https://theappeal.org/the-lab/explainers/risk-assessment-explained ("Nearly every U.S. state and the federal system have implemented risk assessment in some form.").

153. Associated Press, *Policy Group to Expand Research of Pretrial Risk Assessments*, ALBUQUERQUE J. (Apr. 25, 2018, 12:42 PM), https://www.abqjournal.com/1163146/policy-group-pushes-for-risk-assessment-to-score-defendants.html; David G. Robinson & Logan Koepke, *Civil Rights and Pretrial Risk Assessment Instruments*, SAFETY AND JUST. CHALLENGE 3 (2019), http://www.safetyandjusticechallenge.org/wp-content/uploads/2019/12/Robinson-Koepke-Civil-Rights-Critical-Issue-Brief.pdf ("[T]he best available evidence . . . does not clearly establish the impacts of these instruments: Although there is no evidence that they decrease public safety, it remains unclear whether these tools typically cause substantial and lasting reductions in jailing.").

154. Mayson, *supra* note 24, at 2251 ("[P]rediction functions like a mirror. The premise of prediction is that, absent intervention, history will repeat itself. So what prediction does is identify patterns in past data and offer them as projections about future events. If there is racial disparity in the data, there will be racial disparity in prediction too. It is possible to replace one form of disparity with another, but impossible to eliminate it altogether.").

disproportionality of a jurisdiction's detained pretrial population. The evidence of their effectiveness overall is exceedingly thin.[155]

New Jersey's experience is instructive. The state virtually eliminated cash bail in 2014 and overhauled its pretrial justice system entirely, moving to a system focused on measuring and forecasting risk of failure to appear or threat to public safety to guide judges' pretrial detention decisions.[156] Part of that overhaul was the implementation of a pretrial risk assessment instrument, the Public Safety Assessment ("PSA"), developed by the Laura and John Arnold Foundation, now called Arnold Ventures.[157] The PSA examines nine so-called "risk factors" to assess the risk of new criminal activity—specifically new violent criminal activity—along with the likelihood of one's failure to appear pending the resolution of their case.[158] The factors assessed amount to the accused's age at current arrest, criminal history—including prior violent and nonviolent misdemeanor and felony offenses—prior failures to appear, and prior carceral sentences.[159]

New Jersey's turn to risk assessment was made in tandem with a host of other changes to its pretrial system. Among those changes were: a presumption that favors release on nonmonetary conditions over monetary bail; a narrowing of the grounds on which the accused can be detained pretrial; and a requirement that a prosecutor file a detention motion and

155. Stevenson, *supra* note 24, at 341 ("[T]here is a sore lack of research on the impacts of risk assessment in practice. There is no evidence on how the use of risk assessment affects racial disparities. There is no evidence that the adoption of risk assessment has led to dramatic improvements in either incarceration rates or crime without adversely affecting the other margin."). Stevenson's research demonstrated that the implementation of bail reform measures in Kentucky that included the use of a pretrial risk assessment produced limited decarceratory results and no effect on racial disparity:

> [T]he net effects on the overall release rate were small. Furthermore, they were not permanent: the sharp change in practices and outcomes that occurred right after the law was implemented eroded over time as judges returned to their previous bail-setting practices. Within a couple of years, the pretrial release rate was lower than it was before the bill, and lower than the national average. . . . Once county effects were taken into account, racial disparities remain constant throughout the time period of the analysis.

Id. at 309.

156. Lisa W. Foderaro, *New Jersey Alters Its Bail System and Upends Legal Landscape*, N.Y. TIMES (Feb. 6, 2017), https://www.nytimes.com/2017/02/06/nyregion/new-jersey-bail-system.html.

157. *See generally APPR Vision and Mission*, APPR, https://www.psapretrial.org/about/factors (last visited Jan. 8, 2020).

158. *Public Safety Assessment: New Jersey Risk Factor Definitions*, N.J. CTS. 1, 1–4 (Dec. 2018) https://www.njcourts.gov/courts/assets/criminal/psariskfactor.pdf?cacheID=jkRmwcV.

159. *Id.*

overcome—at a hearing—a rebuttable presumption of release by a showing of clear and convincing evidence that detention is warranted.[160]

These changes brought with them significant reductions in New Jersey's pretrial population, leading the state to incarcerate 6,000 fewer people pretrial in 2018 as compared to 2012.[161] That is a noteworthy and commendable reduction. Yet racial disparities in bail decisions persist.[162] According to a 2018 report conducted by New Jersey's Administrative Office of the Courts, Black males continue to be overrepresented in the pretrial incarceration populations, despite the extensive pretrial reforms—and reductions in pretrial incarceration—initiated by the state's overhaul of its criminal legal system.[163] Thus, while the number of Black women who are incarcerated pretrial fell from 44% to 34% from 2012 to 2018, Black men still comprise more than 50% of the state's incarcerated population.[164] And it failed to rectify racial disparities in pretrial detention generally.[165]

The PSA undoubtedly played some role in the reduction of the pretrial population; the presence of simultaneous, significant reforms makes it impossible to measure just how much of a role the PSA played. That is because the PSA is often adopted in conjunction with a host of other pretrial reforms, obscuring what has produced results.[166]

Since the tools have been implemented and expanded rapidly over a short period of time, little data is available to determine their efficacy or fairness.[167] However, even if these data points were readily accessible,

160. ACLU of N.J. et al., *The New Jersey Pretrial Justice Manual*, NAT'L ASSOC. CRIM. DEF. LAWS. 1, 23–30 (Dec. 2016), https://www.nacdl.org/getattachment/50e0c53b-6641-4a79-8b49-c733def39e37/the-new-jersey-pretrial-justice-manual.pdf.

161. Roman Gressier, *Racial Disparities in NJ Bail Persist Despite Reforms: Report*, CRIME REP. (Apr. 4, 2019), https://thecrimereport.org/2019/04/04/racial-disparities-in-nj-bail-persist-despite-reforms-report/.

162. *Id.*

163. Glenn A. Grant, *Report to the Governor and the Legislature*, N.J. CTS. 1, 8 (Apr. 2019), https://njcourts.gov/courts/assets/criminal/2018cjrannual.pdf?c=taP.

164. *See* Gressier, *supra* note 161; Joe Hernandez, *N.J. Officials Finally Release Data on Bail Reform. Their Conclusion? It's Working.*, WHYY (Apr. 2, 2019), https://whyy.org/articles/n-j-officials-have-finally-released-data-on-bail-reform-their-conclusion-its-working. Glenn A. Grant, *Report to the Governor and the Legislature*, N.J. CTS. 1, 21 (2019), https://njcourts.gov/courts/assets/criminal/cjrannualreport2019.pdf?c=NF1 (noting that 55% of the New Jersey jail population is Black and that the "demographic distribution of male inmates is similar to the total population").

165. Gressier, *supra* note 161.

166. *See* SOUTHERLAND, *supra* note 11, at 9 (highlighting remarks from Kristin Bechtel at the Arnold Foundation, which launched the PSA, explaining that the tool "was just one element of a package of changes"); *see also The Impact*, MOVEMENT ALL. PROJECT (last visited June 1, 2021), https://pretrialrisk.com/the-impact/ (describing the limited impact of pretrial risk assessments).

167. Buskey & Wood, *supra* note 145 (highlighting a lack of validation studies and problems with existing validation effort).

pretrial risk assessments—like all actuarial criminal legal system tools—raise a more fundamental concern about decision-making in the criminal legal system. The individual who often matters most after the accused is the person deciding their fate. These tools tend to ignore that decisionmaker.

3. Critique: Looking in All the Wrong Places

The desire to reduce the pretrial detention population and to address unwarranted racial disparities are the oft-stated motivations that animate the introduction and use of algorithmic tools in pretrial decision-making.[168] These goals are commendable. The thinking behind them finds root in optimism: If judges could just choose the right people to detain or set free, we would have a fairer, less biased system.[169] It is also logical. Ultimately, judges are the ones who make the decisions that lead to a robust, racially disparate, predominately poor pretrial population. Unfortunately, there is little evidence to suggest that algorithmic tools, alone and as currently constructed, can meet the laudable goals and optimism that often drives their use.[170] Critiques abound explaining why and how the tools fall short from a practical and civil rights perspective.[171]

---

168. DESMARAIS & LOWDER, *supra* note 8; Sarah Picard et al., *Beyond the Algorithm: Pretrial Reform, Risk Assessment, and Racial Fairness*, CTR. FOR CT. INNOVATION 1, 3–4 (2019), https://www.courtinnovation.org/sites/default/files/media/document/2019/Beyond_The_Algorithm.pdf; Nuefeld, *supra* note 12.

169. Sarah Brayne & Angele Christin, *Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts*, SOC. PROBS., Mar. 2020, at 1, 3, 13; Nuefeld, *supra* note 12; *Bail Reform*, ARNOLD VENTURES, https://www.arnoldventures.org/work/release-decision-making/ (last visited Jan. 13, 2020).

170. *See* Hill, *supra* note 27 (analyzing the risks of algorithmic tools for racial justice); *see also* Ethan Corey, *New Data Suggests Risk Assessment Tools Have Little Impact on Pretrial Incarceration*, APPEAL (Feb. 7, 2020), https://theappeal.org/new-data-suggests-risk-assessment-tools-have-little-impact-on-pretrial-incarceration (describing failures across jurisdictions); Henry, *supra* note 152 (analyzing the pitfalls of risk assessment tools and why organizers oppose using them).

171. For example, twenty-seven researchers, scholars, and advocates signed an open statement of concern—submitted to California and Missouri—regarding the use of actuarial risk assessment as a means of lowering pretrial jail populations. Chelsea Barbaras et al., *Technical Flaws of Pretrial Risk Assessments Raise Grave Concerns*, MIT MEDIA LAB (July 17, 2019), https://damprod.media.mit.edu/x/2019/07/16/TechnicalFlawsOfPretrial_ML%20site.pdf; *see also* Koepke & Robinson, *supra* note 30, at 1750; Mayson, *supra* note 24, at 2227–51 (describing the equality tradeoffs of algorithmic risk assessment instruments). *See* Robinson & Koepke, *supra* note 153, at 3–9 (describing the civil rights concerns raised by algorithmic pretrial risk assessments tools); THE LEADERSHIP CONFERENCE ON CIVIL AND HUMAN RIGHTS, THE USE OF PRETRIAL "RISK ASSESSMENT" INSTRUMENTS: A SHARED STATEMENT OF CIVIL RIGHTS CONCERNS (2018), http://civilrightsdocs.info/pdf/criminal-justice/Pretrial-Risk-Assessment-Full.pdf (describing civil rights concerns raised by algorithmic risk assessments); Julia Angwin et al., *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-

One rarely explored reason is that the tools themselves do little to shape or change the behavior of the actors who are ultimately making decisions in ways that would reduce pretrial detention or confront racism in the criminal legal system. That is because the tools are aimed at the accused rather than the people making decisions about them. They are designed to forecast what those individuals might do in light of their prior history as constructed by the world around them and unique characteristics.[172] Further, the instrument correlates those factors to what others have done in the past. The focus is entirely on the individual before the court. Wholly absent from the frame is what decisionmakers in the system have or have not done in the past when faced with a particular decision point, set of facts, or series of allegations.[173] In other words, there are no risk assessment instruments in use that purport to measure the decision-making of actors within the system by examining the behavior of those actors.

In part, that is because our own biases about systemic reform and the limits of politics and the law have stifled our imagination around points of potential intervention, particularly when it comes to pretrial justice. We understand that racial disparity exists. We concede that our jails hold a racially disparate share of poor people and people of color in pretrial detention.[174] We also presume that if we provide judges enough data about those individuals, they will make fairer, racially just decisions.

Yet that framing ignores a key measure of disparity: the actual behavior of actors in the system. Research has demonstrated that implicit and explicit bias plays a significant role in decision-making throughout the criminal legal system, and in particular in bail determinations.[175] Judges, like anyone else, are subject to biases that shape their decisions.[176] The fact that people of color are treated worse than their white counterparts at every stage of the criminal legal system is not solely a reflection of the behavior of those individuals or indicative of the things they are accused of having done. Rather, that disparity in treatment flows from the biased judgments of

in-criminal-sentencing (describing how an algorithmic risk assessment erroneously overestimated the risk posed by Black people while underestimating the risk posed by their white counterparts).

172. *See* Henry, *supra* note 152 (describing how risk assessment algorithms predict outcomes).

173. *See* Desmarais & Lowder, *supra* note 8 (describing the descriptive factors used most commonly by algorithmic tools, which do not include information about the decisionmaker).

174. Wendy Sawyer & Peter Wagner, *Mass Incarceration: The Whole Pie 2019*, PRISON POL'Y INITIATIVE (Mar. 19, 2019), https://www.prisonpolicy.org/factsheets/pie2019_allimages.pdf.

175. Cynthia E. Jones, *"Give Us Free": Addressing Racial Disparities In Bail Determinations*, 16 N.Y.U. J. LEGIS. & PUB. POL'Y 919, 939–44 (2013).

176. *See* Jerry Kang et al., *Implicit Bias in the Courtroom*, 59 UCLA L. REV. 1124, 1148, 1150–52 (2013) (describing the influence of implicit bias on criminal justice decision-making); Jeffrey J. Rachlinksi et al., *Does Unconscious Racial Bias Affect Trial Judges?*, 84 NOTRE DAME L. REV. 1195, 1208 (2009) (same).

powerful system actors about who poses a danger and who does not, who will likely return and who will not. Thus, even if tools were able to precisely forecast what an individual may do while awaiting disposition of their case, there is no way to ensure that the same biases that shape decision-making now would disappear altogether or cease to play an outsized role in decision-making regardless of the forecast.

Indeed, one study of the adoption and implementation of algorithmic tools in the criminal legal system documented professional resistance by judges and prosecutors to the adoption of tools.[177] The study also found that, far from correcting the biased exercise of discretion, "predictive algorithms in fact displace[] discretion to less visible parts" of the criminal legal system, such that "legal professionals manipulate the data at their disposal to regain the autonomy that they feel is being threatened by the adoption of . . . [new] technologies."[178] Shifts in discretion just lead to "new increases in discriminatory behaviors."[179]

Pretrial risk assessments currently in use track the concerns relevant to statutes governing pretrial release—one's risk of flight or the potential that an individual might be rearrested.[180] Being tethered to statutory considerations at the expense of any other inquiries limits their overall utility. They do not forecast the risk of being wrongfully detained or having bail set too high by a particular judge. They do not tell us whether a prosecutor's office unjustly but consistently seeks detention or bail for those they charge with crimes. Nor have they been used to provide any real insights about judicial behavior. In a regime grounded on evidence-based practices, there is little—if any—inquiry about the evidence that judges (or other criminal legal system actors) are behaving in unbiased ways or imposing pretrial conditions that comport with justice.

Given what we know, the consequences of ignoring the behavior of system actors are significant. First, it ensures that we will continue to focus

177. Brayne & Christin, *supra* note 169, at 7–10; *see also* Stevenson, *supra* note 21, at 341–69 (evaluating how Kentucky judges used risk assessment instruments).

178. Brayne & Christin, *supra* note 169, at 13.

179. *Id.* at 14.

180. Koepke & Robinson, *supra* note 30, at 1752–54. Those concerns include things like one's ties to a community because of stable employment, a steady address, a history of failures to appear in court, prior criminal history, the nature and character of prior criminal convictions, pending or current charges, age, and marital status. Sarah L. Desmarais, et al., *Predictive Validity of Pretrial Risk Instruments: A Systematic Review of the Literature*, J. CRIM. JUST. & BEHAV. (on file with author); *see also Community Justice Exchange, an Organizer's Guide to Confronting Pretrial Risk Assessment Tools in Decarceration Campaigns*, CMTY. JUST. EXCH. 36–41 (Dec. 2019), https://static1.squarespace.com/static/5ba95e4c51f4d408d6784c85/t/5defce44bdfdf024df3b87f8/1575997003585/CJE_PretrialRATGuide_FinalDec2019Version.pdf (detailing variables commonly considered by algorithmic risk assessments).

on those being consumed by the system while failing to scrutinize the system and those who make the decisions that produce harm and burden the accused. Second, a system that fails to engage in critical self-evaluation and corrective behavior undermines any faith that we can put in reform efforts. We rightly expect those convicted of crimes to reflect on their behavior and change it for the better. Our failure to expect the same of system actors undermines the integrity of the system itself.

D. Algorithmic Tools and Sentencing

1. Theory

Over the last two decades, jurisdictions nationwide have adopted algorithmic risk assessment tools to guide sentencing decisions.[181] The shift to the use of these tools to assist sentencing decisions finds root in a larger movement of to engage in evidence-based practices to make the criminal legal system "smart, rather than tough, on crime."[182] Criminal legal systems have embraced these tools largely on the hope that they can distinguish between people who pose a high or low risk of reoffending with greater precision and, in turn, foster a more efficient and effective allocation of limited sentencing resources.[183]

These tools first emerged in the 1920s as guides to assist parole decision-making.[184] Correctional authorities used them to shape the administration of punishment and to help identify the correctional interventions one should receive if incarcerated or under some form of supervision.[185] University of Chicago Professor Ernest Burgess was among the first to develop a risk assessment instrument, designed to predict an individual's likelihood of success on parole based on an examination of

---

181. Brandon L. Garrett, *Evidence-Informed Criminal Justice*, 86 GEO. WASH. L. REV 1490, 1514 (2018) (noting that "an increasing number of states use risk-based instruments to inform decisionmaking at sentencing" and that the use of these tools has been countenanced and encouraged by state supreme courts and statutes.); Erin Collins, *Punishing Risk*, 107 GEO. L.J. 57, 63 (2018) (recounting the growth in use of algorithmic tools at sentencing)

182. Collins, *supra* note 181; *see also* Barkow, *supra* note 10, at 1619 (describing the sentencing guidelines regime as arising out of "dissatisfaction with discretionary and indeterminate sentencing regimes that focused too much on individualization and not enough on avoiding unjust disparities").

183. Dawinder Sidhu, *Moneyball Sentencing*, 56 B.C. L. REV. 671, 673 (2015).

184. Richard A. Berk & Justin Bleich, *Statistical Procedures for Forecasting Criminal Behavior: A Comparative Assessment*, 12 CRIMINOLOGY & PUB. POL'Y 513, 513 (2013).

185. Collins, *supra* note 181, at 79–80.

twenty-one factors.[186] A competing tool, developed by criminologists at Harvard Law School, narrowed the number of predictive factors to seven.[187]

Despite their label, sentencing risk assessments were not intended to determine sentence length.[188] Instead, they generally were:

> [C]reated to guide decisions about how to administer punishment, not about how much punishment is due. In fact, the social scientists who developed the tools that are being incorporated into sentencing decisions expressly disavow their use to "assist in establishing the just penalty," specifically in decisions about whether to incarcerate and the length of the sentence.[189]

Nevertheless, they inform a judge's decision about the length of a person's sentence; they also shape judgments about where an incarceratory sentence will be served, whether the sentence will include supervision, or some form of diversion.[190]

Sentencing risk assessment instruments have become a common feature of the presentence investigation. Presentence authorities—often within the organizational confines of the court system—typically administer the instrument during a presentence investigation and provide the results to the court, defense counsel, the prosecution, and the person facing judgment as a data point to be considered when fashioning an appropriate sentence. These instruments generally seek to forecast one future outcome. They look to quantify the risk that someone will reoffend in some way that undermines public safety.[191] This consideration of future dangerousness and public safety risk at sentencing has been endorsed by the U.S. Supreme Court and has become essential to criminal sentencing.[192]

The development of sentencing risk assessment instruments follows a familiar process. Constructing an algorithmic tool of this sort requires first collecting "data on people charged or convicted of crimes in the past as a base population."[193] Data sources vary, but generally draw from observations of those released from prison or those referred to probation or some other

186. Harcourt, *supra* note 30, at 58; *see also* Berk & Bleich, *supra* note 184, at 513 (citing Professor Burgess's study as one example of predictive tools dating back to the 1920s).

187. Harcourt, *supra* note 30, at 61. The criminologists, professor Sheldon Gleuck and research assistant Eleanor Gleuck, arrived at seven factors to refine their tool to a narrow set of factors, and in turn, fewer predictive variables, guided by their research and data collection. *Id.* at 60–62.

188. Collins, *supra* note 181, at 61.

189. *Id.*

190. *Id.* at 67–71.

191. *Id.*

192. Garrett, *supra* note 181, at 1513–14.

193. Jessica Eaglin, *Constructing Recidivism Risk*, 67 EMORY L.J. 59, 73 (2017).

form of supervision.[194] It also may come from different geographic locations from the venue of the sentencing at issue as well—different regions of a state, the United States, and, in some instances, other countries.[195] Designers then undertake to define recidivism—whether that is an arrest, an arrest plus a formal charge, a final adjudication, or some other conduct.[196] Tool developers then create a statistical model to identify factors that bear a statistically significant correlation with recidivism.[197] That model is the framework for the actuarial risk assessment tool.

Ultimately, the number of factors varies with each instrument, but, generally, they "consider 'static' factors that the [person to be sentenced] can do nothing about (like prior crimes or age) and 'dynamic' risk factors that [may change over time] (like substance abuse or impulsivity)."[198] Most include consideration of four categories of factors: (1) criminal history, (2) antisocial attitude, (3) demographics, and (4) socio-economic status.[199] Like pretrial risk assessments, sentencing risk assessments produce a numerical score by evaluating whether an individual possesses certain risk factors—such as criminal history, socio-economic status, mental health status and history, marital status, and a range of demographic features.[200] That information may be collected through a structured interview with the person to be assessed, by way of a questionnaire to be completed voluntarily by the person to be sentenced, or, in some instances, through publicly accessible data about the individual.[201] The score is associated with a category of recidivism risk—usually low, medium, or high.[202]

The character, nature, and accuracy of the prediction varies with the algorithmic tools used.[203] So too does the level of transparency of the factors considered by the tool and the weight given to them.[204] Thus, there is no standard level of offense or type of recidivism that these tools measure—serious violence or minor criminal behavior may be among the predictive

194. *Id.* at 74.

195. *Id.* at 74–75.

196. *Id.* at 75–76.

197. *Id.* at 78–79.

198. .Christopher Slobogin, *Principles of Risk Assessment: Sentencing and Policing*, 15 OHIO ST. J. CRIM. L. 583, 584–86 (2018).

199. Eaglin, *supra* note 193, at 83.

200. *See* Slobogin, *supra* note 198, at 584–86 (describing three statistically driven risk assessment instruments that are representative of sentencing risk assessments). Tool designers determine "which predictive factors observed in the statistical model" used to construct the algorithmic tool will ultimately be included in the tool. Eaglin, *supra* note 193 at 81–88.

201. Eaglin, *supra* note 193, at 85.

202. Eaglin*, 'supra* note 31, at 92.

203. Slobogin, *supra* note 198, at 587–92.

204. Garrett, *supra* note 181, at 1515.

outputs.[205] Nor is there a standard temporal limit on when reoffense may occur.[206] Some tools address risk management—what is needed to prevent recidivism—while others only produce a recidivism risk forecast.[207] Notwithstanding the fact that "predicting more serious offenses is more challenging than predicting low-risk offenders,"[208] proponents of algorithmic tools at sentencing posit that they regularly outperform human judgments alone.[209] Proponents also claim that the tools will "increase public safety by reducing recidivism. . . . increase[] the accuracy of decisions judges are already making. . . . [and benefit] the public, who save money while avoiding future victimization" and people convicted of crimes who avoid incarceration.[210] A look at the tools in practice tells a different story.

2. Practice

As with algorithmic tools in policing and the pretrial system,[211] algorithmic sentencing tools have not fully delivered the desired results of less biased sentencing or reductions in recidivism.[212] Indeed, a recent empirical study of Virginia's use of algorithmic tools at sentencing provides insights about the wide gulf between the theoretical promise these

205. Slobogin, *supra* note 198, at 587; *see also* Collins, *supra* note 181, at 64–65, 107 (explaining that tools vary in the type of recidivism they predict—from rearrest to conviction, to reconviction for any offense, including a misdemeanor, felony, or violation of court-imposed supervision).

206. Slobogin, *supra* note 198, at 587–88.

207. *Id.* at 588.

208. Garrett, *supra* note 181, at 1515.

209. *Id.* at 1514; *see also* John Monahan, *A Jurisprudence of Risk Assessment: Forecasting Harm Among Prisoners, Predators, and Patients*, 92 VA. L. REV. 391, 408 (2006) (explaining that "[t]he general superiority of actuarial over clinical risk assessment in the behavioral sciences has been known for half a century"). The debate on accuracy of the tools is not over. A 2018 Dartmouth College study found that people responding to an online survey were able to predict risk about as well as the COMPAS risk assessment. Julia Dressel & Hany Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, SCI. ADVANCES, Jan. 17, 2018, at 1, 3; Collins, *supra* note 181, at 95.

210. Collins, *supra* note 181, at 72.

211. *See* Huq, *supra* note 11, at 1074–85 (describing the widespread use of algorithmic and actuarial tools in sentencing and noting that it is "'improbable' that that any convicted felon, whether an adult or juvenile, would be sentenced today without the aid of some sort of actuarial risk instrument, albeit not necessarily one that employs algorithmic means.") (internal citations omitted).

212. *See id.* at 1049, 1052 (explaining that actuarial sentencing fosters incarceration and incapacitation, undermining efforts to curb recidivism through rehabilitation and the provision of services; that even the best instruments are wrong at least 30% of the time; and that fiscal savings are difficult to calculate and often outweighed by the human costs of inaccurate predictions and unnecessary incarceration); Sonja Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803, 806, 842 (2014) (noting that actuarial risk assessment does not provide anything close to a precise prediction of individual risks).

instruments hold and the way they operate in practice.[213] Critically, the study marks "the first evaluation of how risk assessment at sentencing affects outcomes relative to the status quo."[214] As relevant here, one of the Virginia tools studied was used in conjunction with sentencing guidelines to divert low-risk people convicted of nonviolent offenses from jail or prison.[215]

The results of the study encapsulate the difficulties algorithmic tools face in meeting the promise their proponents believe they hold. In short, "Virginia's nonviolent risk assessment reduced neither incarceration nor recidivism; its use disadvantaged a vulnerable group (the young); and failed to reduce racial disparities."[216]

Although sentences for those with high risk scores increased and those with low risk scores decreased, there was scant evidence that the tool yielded a reduction in recidivism.[217] The instrument suggested that judges should have imposed lengthier sentences on young people than were actually being imposed on youth,[218] meaning that if judges followed the tool's recommendations, there would have been an increase in sentences for young people. Nevertheless, the tool did lead to a slightly greater chance of incarceration for young people and an increase in sentence length for youth.[219]

Racial disparities in sentencing were largely unchanged by the tool, though Black people scored substantially higher—and therefore riskier—than their white counterparts.[220] Racial disparities grew in courts where the tool was viewed as the most influential, largely due to the tendency of judges to exercise more leniency for white people with high risk scores than for Black people with high risk scores.[221]

This study also shed light on the role of discretion by judges when given an algorithmic tool. Among the findings were that judges were three percentage points less likely to divert Black people in the highest risk

---

213. Megan T. Stevenson & Jennifer L. Doleac, *Algorithmic Risk Assessment in the Hands of Humans* 1, 5 (Nov. 18, 2019) (unpublished manuscript), https://ssrn.com/abstract=3489440.

214. *Id.* at 5.

215. *Id.* at 2.

216. *Id.* at 5. The study's authors provide a number of possible explanations for this set of results. Among those explanations are the exercise of discretion by sentencing judges whose decisions are shaped by a host of factors; judges gaining familiarity with the forecasts that a risk instrument produces; the willingness of a judge to consult an algorithmic tool; and the way judges make use of the information conveyed by the algorithmic tool. *Id.* at 22–29

217. *Id.* at 2.

218. *Id.*

219. *Id. at* 3.

220. *Id.* at 2.

221. *Id.*

category out of the formal system than white people in the same category.[222] Judges also chose whether to follow or deviate from the algorithmic tool when factors like race, gender, or socioeconomic status were at play—factors that shaped the judge's view of the person before the court and the circumstances that led them into the criminal legal system.[223] Judicial discretion actually minimized the significant increase in the chances of incarceration for a young person that would have resulted from a faithful adherence to the instrument's forecast.[224] At the same time, judges who used the algorithmic tools the most were also more likely to be more lenient to white people with high risk scores than they were with Black people who similarly scored high risk.[225]

These real-world consequences highlight the challenges that come with the development and implementation of algorithmic tools. They also underscore the very real difficulty of forecasting an individual's future based on what we know about other individuals. More to the point, they underscore the shortcomings of profiling.

3. Critique: Racial Profiling 2.0?

By nature, algorithmic tools produce their risk scores by analyzing group-level data that correlates with certain types of behavior of interest to a decisionmaker.[226] The tools then assign a score that approximates the relationship between the characteristics possessed by the group and the behavior engaged in by members of the group.[227] The similarity between the individual being assessed and the group from which the data is drawn produces a forecast of what an individual may do. In other words, the tools "ascribe a blanket risk profile to all individuals in a group," recommending treatment based on an individual's association with a group.[228] Thus, the tools rank people convicted of crimes "according to likelihood of engaging in criminal behavior based on the behavior of the individuals in the

222. *Id.* at 25.

223. *Id.* at 26.

224. *Id.* at 27.

225. *Id*. at 29. Another study of Virginia's Nonviolent Risk Assessment ("NVRA") revealed an additional concern. The tool was developed with the stated goal of identifying people convicted of nonviolent crimes at the lowest risk of recidivism for diversion from prison. Brandon Garrett & John Monahan, *Assessing Risk: The Use of Risk Assessment in Sentencing*, 103 JUDICATURE, SUMMER 2019, at 42, 45 (2019). A review of sentencing data from 2016 concerning the use of the NVRA by judges in diversion decisions revealed that "many—indeed, most—defendants eligible for [ ] alternative sentences did not receive them." *Id.*

226. *See* Nicholas, *supra* note 11 and accompanying text.

227. *See* Eaglin, *supra* note 193, at 85–88 (describing how sentencing algorithmic tools are designed and constructed).

228. Sidhu, *supra* note 183, at 702.

underlying data set."[229] In that sense, actuarial risk assessments operate as a form of digital profiling, prescribing the treatment of an individual based on their similarity to, or membership in, a group. Forecasts based on actuarial data provide us with insights about groups of people but reveal far less about individuals.[230] When it comes to sentencing, the tools become a way of asking whether the person before the court is more, less, or equally dangerous as a group of people based solely on the statistical similarities between the group and the individual.[231]

Such an approach is troubling, to say the least. Treating someone in a specific way because they share the characteristics of a group is the essence of profiling.[232] Such conduct offends an axiomatic principle that cuts across the criminal legal system and bears particular significance at sentencing: individuals should be treated as individuals, not based on their membership in, or shared characteristics with, a particular group.[233] Put differently, "our criminal law punishes people for what they do, not who they are."[234] That edict carries even more weight when one considers the fact that sentencing risk assessments, by potentially suggesting a lengthier term of incarceration based on a rough forecast that one may recidivate, are in essence punishing individuals not only for crimes they have not yet committed but for anything they may ever do at any point in the future.

Since actuarial sentencing takes root in the toxic soil of profiling and encourages the analysis of characteristics that correlate with recidivism, it necessarily drives judges to consider factors that may have nothing to do with culpability.[235] Actuarial risk assessments not only "incorporate a range of non-culpable characteristics into their calculations, most of [them]

229. Eaglin, *supra* note 193, at 85.

230. Brian Netter, *Using Group Statistics to Sentence Individual Criminals: An Ethical and Statistical Critique of the Virginia Risk Assessment Program*, 97 J. CRIM. L. & CRIMINOLOGY 699, 714 (2007) (explaining that "[s]tatistical predictions speak to group tendencies, not individual proclivities")

231. Sidhu, *supra* note 183, at 675.

232. *Profiling*, BLACK'S LAW DICTIONARY (11th ed. 2019) ("The practice or an instance of using particular information about someone or something to infer other probable characteristic.").

233. Sidhu, *supra* note 183, at 702–03 (explaining that punishment is to be inflicted because of one's conduct, not group membership). The Supreme Court surfaced this concern in *Batson v. Kentucky*, which set forth a framework to address racial discrimination in jury selection. 476 U.S. 79 (1986). The Court explained that its concern rested on the notion that prosecutors were making decisions about who to strike from juries based on their race and alleged resultant affinity for a particular group, rather than on an individualized consideration of their fitness as a juror. *Id.* at 87–88.

234. *Buck v. Davis*, 137 S. Ct. 759, 766 (2017).

235. Collins, *supra* note 181, at 103–04; *see also* Eaglin, *supra* note 31, at 99 (explaining that predictive analytics weighs factors that fall outside an individual's control, exposing "[t]ensions regarding what counts at sentencing and the meaning of fairness.").

omit . . . the crime for which the [convicted person] is being punished."[236] Thus, the risk assessment suggests a punishment that does not reflect a consideration of the crime of conviction, but instead relies on factors such as one's gender, education, employment history, and mental health status.[237]

This raises yet another profiling-related concern. It is not hard to imagine a host of other factors that are deemed relevant to sentencing through the lens of an actuarial risk assessment because they are correlated with recidivism.[238] Such factors may also be associated with distinct disadvantages faced by communities of color.[239] For example, imagine a sentencing risk assessment that considers one's zip code, level of education, marital status, familial ties, and parental criminality. Given the way structural inequality influences life outcomes along racial lines, all of these factors unfairly disadvantage Black people facing sentencing.[240]

That is problematic for at least two reasons. First, it perpetuates racially disparate treatment at sentencing. People with what are considered negative characteristics will be viewed as recidivism risks and will therefore warrant harsher treatment. If those people happen to be Black, racially disparate treatment will be the result. Second, it forces those who rely on risk assessment to equate correlation with causation. In doing so, decisionmakers must forgo consideration of context and nuance—the reasons why the individual before the court may be different from all those who previously appeared for sentencing. The result is a sentencing regime that either punishes a person or dispenses mercy based on who they are in comparison to others, rather than what brought the individual before the court, what they did, and who they might become in the future with or without the intervention of a criminal sanction.[241]

---

236. Collins, *supra* note 181, at 103. One example of incorporating nonculpable characteristics is the consideration of marital status as it relates to recidivism. Marital status may connote less time spent outside of the house, which is the true predictor of recidivism. If an instrument only considers marital status, but does not consider time outside of the home, the use of the correlated variable (marital status) in the instrument instead of the true variable (time outside the house), means that those single people who do not spend time outside the house will be scored riskier because they are not married, even though as an individual, they may be less risky. Netter, *supra* note 230, at 715.

237. Collins, *supra* note 181, at 104–05. This is a variation on the flawed data as destiny, garbage in, garbage out critique detailed. *See supra* Section I.B.3.

238. Sidhu, *supra* note 183, at 702.

239. Eaglin, *supra* note 31, at 95–97.

240. *Id.* at 96–97.

241. Collins, *supra* note 181, at 107 (noting that those who benefit from actuarial sentencing benefit because of "'' their "relative privilege" in the form of "access to educational and employment opportunities, [and] a low-crime zip code. . . ."); *see also* Sidhu, *supra* note 183, at 707–10 (explaining that risk assessments "demand punishment for a group identity over which the individual has no meaningful control").

To be sure, actors throughout the criminal legal system use anecdotal, qualitative, or quantitative data about groups to make judgments about individuals. The routine nature of the practice does not make it less troubling. Saddling judgments about individuals with the behaviors and actions of others who may be similarly situated by age at first arrest, marital status, employment status, or their prior involvement with the criminal legal system raises concerns about equity and justice that are unique when one's freedom is on the line.[242] Most would agree that a just criminal legal system requires those sitting in judgment of the accused to undertake a holistic consideration of the person before them, weighing factors for which an algorithm may not account.[243] Judging people based on their associations with data points flies in the face of the notion of an individualized evaluation of the person standing before the court.

The problem with profiling is highlighted both by the robust debate over the differing measures of fairness of algorithmic risk assessment instruments and the impact that such judgments can have on individuals. In 2016, the news organization *ProPublica* investigated the accuracy of risk assessment scores used in pretrial decision-making in Broward County, Florida.[244] They examined the risk scores of more than 7,000 Broward County arrestees from 2013 and 2014 to evaluate how many arrested people would be charged with new crimes over the next two years.[245] What their investigation uncovered was nothing short of breathtaking. Unreliable forecasts of violent crime were the instrument's hallmark: only 20% of those predicted to commit violent crimes went on to do so.[246] The faulty forecasts not only carried serious racial disparities but also inaccurate predictions of who posed a risk of future criminality. Black people were falsely labeled as future criminals at nearly twice the rate of their white counterparts, while white people were mislabeled as low risk more often than their Black counterparts.[247]

ProPublica foreclosed the possibility that these disparities could result from prior criminal history, age, and gender. Even after controlling for those variables, "Black defendants were still 77% more likely to be pegged as at

242. Netter, *supra* note 230, at 714.

243. *See* Caryn Devins et al., *The Law and Big Data*, 27 CORNELL J.L. & PUB. POL'Y 357, 396 (2017) (explaining that just sentencing requires judicial discretion to "consider the individual holistically, to weigh the competing purposes of sentencing, and to consider factors not accounted for by the Guidelines. In other words, the "frame" of sentencing determinations is fluid and requires case-by-case evaluations. The variables that were important in one sentencing proceeding may be less influential in another. These types of discretionary determinations are inherently not reducible to rigid criteria or models.").

244. Angwin et al., *supra* note 156.

245. *Id.*

246. *Id.*

247. *Id.*

higher risk of committing a future violent crime and 45% more likely to be predicted to commit a future crime of any kind."[248]

Northpointe, the company responsible for the risk assessment instrument that produced those scores, rejected *ProPublica*'s analysis.[249] In doing so, Northpointe argued that the tools they constructed were racially neutral because Black and white people who were labeled high risk were rearrested at the same rates.[250] Thus, Northpointe claimed, the tool accurately sorted individuals without regard to race.[251]

The debate between *ProPublica* and Northpointe[252] raises a point about measuring fairness and equity that illuminates the profiling concern.[253] Each entity is examining notions of fairness and equality through a different set of lenses. For *ProPublica*'s part, the measure of fairness that matters most is error rate balance. Under that rubric, the fairness of the tool depends on preventing any single group or individual from bearing the burden of the mistakes made by the risk assessment instrument.[254] From Northpointe's perspective, the fact that when an individual is labeled high risk, they are

248. *Id.*

249. WILLIAM DIETERICH ET AL., COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity, NORTHPOINTE RESEARCH DEPT. 1, 2 (July 8, 2016), https://www.documentcloud.org/documents/2998391-ProPublica-Commentary-Final-070616.html.

250. *Id.* at 2–3.

251. *Id.* at 11–13; Eaglin, *supra* note 193, at 97.

252. *See, e.g.*, Laurel Eckhouse et al., *Layers of Bias: A Unified Approach for Understanding Problems with Risk Assessment*, 46 CRIM. JUST. & BEHAV., 185, 190–93 (2019) (detailing ProPublica and Northpointe's positions on the COMPAS tool); Emily Berman, *A Government of Laws and Not of Machines*, 98 B.U. L. Rev. 1277, 1328 (2018) (same); Eaglin, *supra* note 193, at 96–97 (same); Anne L. Washington, *How to Argue with an Algorithm: Lessons from the COMPAS-ProPublica Debate*, 17 COLO. TECH. L.J. 131, 148-151 (2018) (same).

253. For an interactive exercise that demonstrates the challenges presented by the fairness tradeoffs at the heart of this debate, *see* Karen Hao & Jonathan Stray, *Can You Make AI Fairer Than a Judge? Play Our Courtroom Algorithm Game*, MIT TECH. REV. (Oct. 17, 2019), https://www.technologyreview.com/s/613508/ai-fairer-than-judge-criminal-risk-assessment-algorithm/. Research has revealed that it is mathematically impossible to simultaneously meet both definitions of fairness in the *ProPublica* and Northpointe debate when the input data that captures the behavior of different groups differs. *See* Alexandra Chouldechova, *Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments*, 5 BIG DATA 153 (2017) ("[A]n instrument that satisfies predictive parity cannot have equal false positive and negative rates across groups when the recidivism prevalence differs across those groups."); Jon Kleinberg et al., *Inherent Trade-Offs in the Fair Determination of Risk Scores*, 2017 PROC. INNOVATIONS THEORETICAL COMP. SCI. 1, 17 (2017), https://arxiv.org/pdf/1609.05807.pdf (describing the tradeoffs of measures of fairness).

254. In this way, the *ProPublica* measure most clearly reflects the concern that may arise when someone is profiled—that they suffer because of what others have done in the past.

rearrested at the same rates as others who share that label—predictive parity—is indicative of the tool's accuracy.[255]

One measure is concerned with mislabeling individuals; the other is concerned that all those who are labeled alike are treated alike. But both measures still look at the behavior of unrelated groups to determine how individuals should be treated: they "evaluate[ ] [a person's] risk using data about other people."[256] That is, no matter how you measure it, the very essence of profiling.[257]

Profiling has consequences. Among the most disturbing is the "ratchet effect."[258] This concept describes a type of feedback loop that produces disparities between groups who come into contact with the criminal legal system repeatedly over time. It is what happens when, for example, police focus law enforcement resources on people who match the profile of those who are incarcerated for a particular criminal activity rather than on those who are actually engaged in that criminal activity. The ratchet effect comes into play because those who match the profile are subject to greater law enforcement attention and scrutiny. That attention leads to more arrests of the profiled group—a type of self-fulfilling prophecy that encourages further profiling and law enforcement focus on those who match the profile. Meanwhile, those who do not match the profile, but are still engaged in criminal activity, do not receive the same level of law enforcement attention, if they receive any at all. Ultimately, the ratchet effect creates the false impression that the only people who commit crimes are those who match a profile.[259] Much like the potential feedback loop forged by predictive policing, actuarial tools at sentencing encourage us to continue incarcerating the same populations repeatedly which, in turn, fosters the inequity that feeds mass incarceration and criminalization.

Thus, by profiling members of the group who are most likely to be rearrested—those deemed high risk who may be "unattached, unemployed, or unskilled"—the system ensures that those individuals are more likely to be jailed, exacerbating the very risk those individuals allegedly pose by placing one more barrier—a term of incarceration—in their way.[260] In light

255. In other words, "[f]airness could be defined as treating everyone the same or it could be defined as giving everyone similar outcomes." Washington, *supra* note 252, at 150. Accordingly, "[t]he central complication is that there is no single measure of racial equality in risk assessment. Instead, there are many possible measures and, in most circumstances, it is impossible to achieve racial equality according to every measure at once." Mayson, *supra* note 24, at 2233.

256. Eckhouse, et al., *supra* note 252, at 198.

257. *See supra* note 232

258. Harcourt, *supra* note 30, at 220 (emphasis omitted).

259. *Id.* at 3.

260. *Id.* at 220.

of the foregoing, it is not hard to see that deep, troubling problems flow from the profiling problem of actuarial risk assessments at sentencing.

II. PROPOSED SOLUTIONS

To this point, I have grappled with the history, design, and implementation of algorithmic tools at three distinct decision points in the criminal legal system—policing, bail, and sentencing. In doing so, I have catalogued the types of problems that accompany the use of those tools. They rest on data infected by racial bias, and therefore produce forecasts that reflect that bias. They are aimed at the people already targeted by the criminal legal system rather than the system or its decisionmakers. And they encourage profiling by recommending a criminal legal system response based on a person's association with a group.

Solutions to these problems do not come easy. The most straightforward would be ending the use of algorithmic tools in the criminal legal system altogether. At first glance, that is a simple fix. However, given the widespread nature of algorithmic tools,[261] it is unlikely to happen any time in the near future. Even if abolition of the tools merits consideration as an ultimate goal, that road will be paved with paradigmatic shifts in the way systems operate. The rapid development and expansion of algorithmic tools can be viewed as providing opportunities to shape those shifts of the system and implement potential solutions. A blunt end to the use of algorithmic tools also fails to account for the nuance and complexity of the problems they present. All tools—and all stages of the system—are not equal. Nor do they distribute their harms evenly. In keeping with that view, what follows is an exploration of the ways that we might mitigate the potential and realized harms that flow from the use of algorithmic tools in the criminal legal system, with an eye toward abolitionist, transformative ends.

A. A Framework for Confronting Algorithmic Tools

A framework to confront the challenges raised by algorithmic tools requires that we interrogate the role of race, its relationship to power, and the influence of both phenomena on the law.[262] If we understand algorithmic

261. Henry, *supra* note 152 ("Nearly every U.S. state and the federal system have implemented risk assessment in some form."); Huq, *supra* note 11, at 1052 (noting that algorithmic tools are "likely to soon become pervasive" in the criminal legal system).

262. KIMBERLÉ CRENSHAW ET AL., PART FIVE: THE SEARCH FOR AN OPPOSITIONAL VOICE, CRITICAL RACE THEORY: THE KEY WRITINGS THAT FORMED THE MOVEMENT xiii (Kimberlé Crenshaw et al. eds., 1995) (noting that goal of critical race theory is "not merely to understand the vexed bond between law and racial power but to *change* it"); Mari J. Matsuda, *Voices of America: Accent, Antidiscrimination Law, and Jurisprudence for the Last Reconstruction*, 100

tools as instruments that carry the potential to reproduce the racial inequity of our criminal legal system, a lens that is rooted in critical race theory and which focuses on and scrutinizes the nature of racial inequality seems not only appropriate but required.[263] A focus on the role of race in shaping the law—and by extension the world that the law inhabits, defines, and regulates—holds the most promise for a fundamental shift in the way algorithmic tools and the American criminal legal system operate.[264]

The next Section applies a racial justice lens to the challenges presented by algorithmic tools. It addresses what these tools would look like and how they would be deployed if we accepted that racism is a permanent fixture; that the exercise of classification parallels the construction of race; that bold changes are needed to combat the reform/retrenchment paradigm and the tendency of the law to favor the status quo; that the voices of the marginalized are the voices that matter; and that we should engage with the nuance and complexity that shapes one's identity. What follows is a discussion about the policy choices that we need to make regarding the balance of power and algorithmic tools in a way that confronts the racism and unfairness that pervades the criminal legal system.[265]

YALE L.J. 1329, 1331 n.7 (1991) (explaining that critical race theory works "to develop a jurisprudence that accounts for the role of racism in American law and that works toward the elimination of racism as part of a larger goal of eliminating all forms of subordination")

263. A racial justice lens challenges the dominant bases for American antidiscrimination law—the notion that colorblindness produces race neutrality and that color consciousness produces racial preferences. Devon W. Carbado, *Critical What What?*, 43 CONN. L. REV. 1593, 1609 (2011). It does so by demonstrating how "'colorblind' laws often serve to further insider privileges along the lines of race, gender, and class, while marginalizing and obscuring social, political, and economic inequality." I. Bennett Capers, *Afrofuturism, Critical Race Theory, and Policing in the Year 2044*, 94 N.Y.U. L. REV. 1, 24–25 (2019). In this way, a racial justice lens "embraces color consciousness . . . as the way to rectify today's racist legal legacies." Adrien Katherine Wing, *Space Traders for the Twenty-First Century*, 11 BERKELEY J. AFR.-AM. L. & POL'Y 49, 51 (2009). These contributions speak to racism's position not as the product of individual biases alone, but as a structural and institutional phenomenon. Carbado, *supra* note 263, at 1612.

264. DOROTHY A. BROWN, CRITICAL RACE THEORY: CASES, MATERIALS, AND PROBLEMS 1 (2014) ("Critical Race Theory askes the question: 'what does race have to do with it?'"). Critical race theory has also been described as a discipline that:

> [E]xhumes the atrocities of our historical past and confronts their continuing curse; it articulates the ways in which race, gender, and class inequality converge and interpenetrate; and it focuses our attention on the problems of structural discrimination, unequal treatment, and the incomplete nature of democracy in our social order.

Michael Omi & Howard Winant, *The Unfinished Business of Race*, *in* RACE LAW STORIES ix (Rachel F. Moran and Devon W. Carbado eds. 2008). For an overview of critical race theory, see Osagie K. Obasogie, *Foreword: Critical Race Theory and Empirical Methods*, 3 UC IRVINE L. REV. 183, 184 (2013); Capers, *supra* note 264, at 20–30; Adrien K. Wing, *Is There a Future for Critical Race Theory?*, 66 J. LEGAL EDUC. 44, 47–53 (2016).

265. Although the focus of this article is on the criminal legal system, the problems of algorithmic tools, and racial justice, the solutions I have suggested could apply with equal force in

B. Accepting the Truth: The Permanence of Racism

Racism is a permanent, fixed feature of American society.[266] It is "constitutive of, rather than oppositional to, American democracy,"[267] and woven into our nation's fabric.[268] It is "an integral, permanent, and indestructible component of this society."[269] This conclusion stems not from a sense of hopelessness or an acceptance of the second-class citizenship and inequity that racism breeds. It instead emanates from a deep, reflective, and clear-eyed examination of America's history and current condition.

Countless scholars have documented, with excruciating detail, the defect that marred America's birth and continues to shape its life: the ideology of white supremacy, which defined superiority and inferiority along racial lines.[270] Indeed, this racist ideology was America's birthright, baked into the country's DNA.[271] It has been with us for at least four centuries.[272] In that time, it has served a number of purposes. It was used to prop up and justify the enslavement of African people in America.[273] It delineated freedom. It was the handmaiden to the criminal legal system. And it is so interwoven within the range of institutions that govern American life that its presence is ubiquitous today.[274] The deep-rooted nature of institutional, structural, and interpersonal racism, when weighed against the current pace of racial justice-oriented reform, leaves little room for us to hope that we can disentangle racism from the American way of life.

The endemic nature of racism bears the weight of a fundamental truth worthy of acceptance. Yet doing so—actually accepting the deep-seated nature of racism—presents challenges for those who seek to deploy

---

other domains as well where algorithmic tools are used to sort, identify, and produce forecasts about people or places and where those tools rely on existing data to do so.

266. Carbado, *supra* note 263, at 1613.

267. *Id.* at 1613; Wing, *supra* note 264, at 48

268. ANTHONY C. THOMPSON ET AL., A PERILOUS PATH: TALKING RACE, INEQUALITY, AND THE LAW 2 (EDS. 2018).

269. DERRICK BELL, FACES AT THE BOTTOM OF THE WELL: THE PERMANENCE OF RACISM, ix (1992).

270. *See generally*, Nikole Hannah-Jones, *Our Democracy's Founding Ideals Were False When They Were Written. Black American Have Fought to Make Them True.*, N.Y. TIMES MAG. (Aug. 14, 2019), https://www.nytimes.com/interactive/2019/08/14/magazine/black-history-american-democracy.html; THOMPSON ET. AL., *supra* note 268, at 88; IBRAM X. KENDI, STAMPED FROM THE BEGINNING: THE DEFINITIVE HISTORY OF RACIST IDEAS IN AMERICA (2016).

271. Hannah-Jones, *supra* note 270.

272. *Id.*

273. ALEXANDER, *supra* note 3, at 28–31.

274. *Id.* at 26–30; Bryan Stevenson, *Why American Prisons Owe Their Cruelty to Slavery*, N.Y TIMES MAG. (Aug. 14, 2019), https://www.nytimes.com/interactive/2019/08/14/magazine/prison-industrial-complex-slavery-racism.html.

algorithmic tools as a means to attack decision-making infected by implicit and explicit racism. One of the more difficult challenges of accepting the omnipresence of racism is the natural disappointment that comes with realizing that there is no way to eradicate it. Yet that realization obscures what should be the target of our efforts when we seek to employ algorithmic tools in the criminal legal system. Rather than attempting to solve racism, acceptance that racism is a permanent, fixed feature forces us to confront and take stock of the role that racism plays as we design, implement, and engage in oversight of algorithmic tools. That is true not only in the data upon which the tools rely but in the targets at which those tools are leveled, the outputs that those tools produce, and the very institution in which the tools are deployed. The policy recommendations in this Section are informed by, and flow from, recognition and acceptance of this basic premise, with good reason. Our times demand it, and our reality dictates it. Not only because the ideology of racial supremacy and inferiority has shaped American society and its governing institutions, but because the data tells us that the same ideology casts an inescapable shadow over policy and practice in the criminal legal system today.[275]

We cannot hope to change the current state of affairs if we proceed as though the status quo is divorced from our history and our reality. An intentional and focused orientation toward that history and a fulsome response to what it has produced is necessary. This is not a radical idea. It is a suggestion that we exchange those values that blind us to our past for those which acknowledge that history and work to address it—something akin to what we might call a form of digital reparations.[276]

275. Balko, *supra* note 3.

276. The contrast between the values that blind us to our history and those that require we acknowledge it was most readily illustrated in *Schuette v. Coalition to Defend Affirmative Action, Integration and Immigrant Rights and Fight for Equality By Any Means Necessary (BAMN)*, 572 U.S. 291 (2014). The majority in *Schuette* upheld an amendment to Michigan's constitution that barred race conscious admissions policies in higher education. *Id.* at 315. Justice Sotomayor, in dissent, explained, "My colleagues are of the view that we should leave race out of the picture entirely and let the voters sort it out. . . . We have seen this reasoning before." *Id.* at 380 (Sotomayor, J., dissenting). *See Parents Involved in Comm. Schs. v. Seattle Sch. Dist. No. 1*, 551 U.S. 701, 748 (2007) ("The way to stop discrimination on the basis of race is to stop discriminating on the basis of race."). It is, unfortunately, a sentiment out of touch with reality, one not required by our Constitution, and one that has properly been rejected as 'not sufficient' to resolve cases of this nature. *Schuette*, 572 U.S. at 380 (Sotomayor, J., dissenting). Justice Sotomayor further remarked:

> This refusal to accept the stark reality that race matters is regrettable. The way to stop discrimination on the basis of race is to speak openly and candidly on the subject of race, and to apply the Constitution with eyes open to the unfortunate effects of centuries of racial discrimination. . . . [W]e ought not sit back and wish away, rather than confront, the racial inequality that exists in our society. It is this view that works harm, by

C. The Implications of Acceptance

Several policy prescriptions emerge when we take seriously the implications of the worldview that racial inequality is not a passing phenomenon but instead a permanent feature. They fall into three categories. First, there are those that relate to the input data fed into the algorithmic tools and adjustments to the forecasts that the tools produce. That means accounting for race in the data on which algorithmic tools rely (the inputs) and in the forecasts (the outputs) that they produce.

Second, there are those that center on the actors who design and use algorithmic tools. In this case, that means requiring actors that seek to use algorithmic tools to detect and remedy the real and potential harms of those tools. It also means placing algorithmic tools in the hands of communities so that they may deploy them to scrutinize system actors.

Finally, there are broader policy prescriptions about the criminal legal system as a whole that can inform when and whether these instruments are useful. This intervention requires countering the turn to raw numbers with attention to the stories of those enmeshed in the criminal legal system, privileging qualitative information over quantitative data.

Ultimately these interventions would serve as a paradigmatic shift in the way the current system operates, opening up the potential for a different criminal legal system. I address each policy prescription in turn, beginning with changes to input data and the forecasts produced by the tools.

1. Accounting for Race in the Inputs

This first category of measures responsive to the permanence of racism requires that we develop tools that credibly account for racism and the disparities it produces, in the same way factors like prior criminal history, employment status, and education are part of the data analyzed by an algorithmic tool. Fully acknowledging the feature-level nature of race in this way means orienting our work to meet the challenge posed by quantifying the role of race and adjust policy accordingly. To some, that may sound like a radical intervention. In reality, it is what justice, in light of history, requires.

Those who accept the reality of racism and develop algorithmic tools could be explicit about the racial dimensions of the inputs. Rather than engaging in the Sisyphean task of attempting to scrub the data of racism, tool designers could attempt to measure the ways racism shapes the data and then account for it in the algorithms they build and the instruments they create.

> perpetuating the facile notion that what makes race matter is acknowledging the simple truth that race *does* matter.

Id. at 381.

The criminal legal system constantly generates data that could be considered as part of such an effort. For example, if a particular precinct engages in discriminatory policing, the disparities that result from those discriminatory practices would be accounted for in the data set, the algorithm, and the outputs.[277] Data from that precinct could be included but discounted by a quantifiable factor because of the racially disparate impact of policing, or weighted by what one might expect to see in the absence of discriminatory policing. Arrest data produced by officers could be quantitatively evaluated and adjusted to reflect policing patterns and behaviors that are otherwise problematic.[278] Predictions that flow from tools that rely on such data might be accompanied by an explicit disclaimer that the data relied upon is tainted by a history of racially discriminatory policing practices.[279]

Racial disparities in areas such as housing, education, health, wealth, employment, and criminal legal system contact are not unknowable or unknown; they are simply ignored or elided. And yet they are fully baked into the outputs of tools which rely on such data points. If we know that to be the case, it is incumbent on us to take stock of that fact. Quantifying how these disparities shape the lives and the experiences of communities, and then discounting the data points by that numerical value is another way of surfacing and accounting for race.[280] One might attempt to quantify a world that we seek, where all races were treated equally by the criminal legal system, and choose to use that data as part of the analysis.[281]

Leveling the algorithmic playing field is undoubtedly a complex and challenging undertaking.[282] Racial inequality can shape institutions and individual lives in ways that can be impossible to quantify. Since race is a construct, and the dimensions of racism transform over time as the political, legal, and social context change, it may not be possible to design a specific

277. Richardson et al., *supra* note 71, at 24–26.

278. *Id.* at 17–20

279. *Id.*

280. Such an accounting is reminiscent of a racial attrition index, which Professor Derrick Bell imagined would be "prepared by social scientists and computer-oriented statisticians [to] provide a dramatic rendering of our social progress and decline." CARVING OUT A HUMANITY 155 (Janet Dewart Bell & Vincent M. Southerland eds., 2020).

281. Some researchers have attempted to account for racial inequality—and the benefit that white people receive because of their race—by proposing that an algorithmic tool be trained on white people alone as the more privileged group. Richard Berk & Ayya A. Elzarka, *Almost Politically Acceptable Criminal Justice Risk Assessment*, 1, 10 (Dec. 31, 2019), https://www.cis.upenn.edu/~mkearns/teaching/ScienceDataEthics/AlmostPC.pdf. Such a solution is by no means adequate, or even necessarily advisable, but serves as an example of the type of work that could be undertaken to address the racial inequality embedded in data.

282. *See* Mayson, *supra* note 24, at 2265–67 (examining the potential in allowing an "algorithm to assess . . . risk factors *contingent on* race" and describing the possible trade-off in predictive ability of the tool).

measure to capture its effects.[283] It may be that we can never account for all the ways bias, inequity, and racism shape people's lives.[284] But that challenge cannot be dispositive.

The criminal legal system often deals in nuance: the appropriate quantum of punishment, the justifications necessary to vindicate law enforcement intrusion, the decision to proceed to trial or plead guilty, and the credibility of a witness at trial as weighed against biases—explicit and implicit—that shape their testimony. Rather than turn away from the complexity, by design the legal system regularly imposes a requirement that decisionmakers confront and consider it, even if that consideration is less than ideal. A ready example is the instruction given to jurors when evaluating the credibility of a witness against the biases that may shape the witness's testimony.[285] While racism may be a permanent force, its permanence does not prevent us from taking stock of its effects, shaping and remaking the tools that guide decisions with close to full knowledge regarding its effects.

2. Accounting for Race in the Outputs

The challenges of quantifying the impact of racial inequality with precision also do not prevent us from having a different set of responses to the data, or the tools that analyze it. If our aim ultimately is to eliminate

283. Carbado, *supra* note 263, at 1611. Race has no biological significance; it only contains the meaning that we give it. *Id.* at 24. "The anthropologist Ashley Montagu was among the first to argue that race is a human invention, a social construct, not a biological one . . . ." ISABEL WILKERSON, CASTE: THE ORIGINS OF OUR DISCONTENTS 24 (2020). The law differentiates between races and determines the racial categories into which we sort individuals and assigns meanings to those categories—both good and bad—in service of a hierarchy that serves the interests of those in power—the status quo. *Id.*

284. To be clear, there is no singular experience or set of unifying characteristics tied to identity. Capers, *supra* note 263, at 25–26; Devon Carbado & Cheryl I. Harris, *Intersectionality at 30: Mapping the Margins of Anti-Essentialism, Intersectionality, and Dominance Theory*, 132 HARV. L. REV. 2193, 2205 (2019). Although the focus here is on race, an ideal approach is one that accounts for multiple, complex grounds of identity that drive oppression, marginalization, and treatment. Kimberlé Williams Crenshaw, *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*, 1989 U. CHI. LEGAL F. 139, 166-67 (1989).

285. For example, New York jurors are instructed as follows:

> As judges of the facts, you alone determine the truthfulness and accuracy of the testimony of each witness. You must decide whether a witness told the truth and was accurate, or instead, testified falsely or was mistaken. You must also decide what importance to give to the testimony you accept as truthful and accurate."

N.Y STATE UNIFIED COURT SYSTEM, *Credibility of a Witness*, CRIMINAL JURY INSTRUCTIONS: INSTRUCTIONS OF GENERAL APPLICABILITY 1, 2 (2018), https://www.nycourts.gov/judges/cji/1-General/CJI2d.Credibility.pdf. While the instructions go on to state that "[t]here is no particular formula for evaluating the truthfulness and accuracy of another person's statements or testimony[,]" jurors are told to consider, among other things, whether the witness harbored "a bias, hostility or some other attitude that affected the truthfulness of the witness's testimony." *Id.*

unwarranted, race-driven disparate treatment in the administration of criminal law, we can readily implement policy responses that help us to achieve that goal. What I am suggesting is a complete repurposing of algorithmic tools for ends that boldly attack the manifestations of racial inequality at a structural and institutional level. In this way, the tools could function like a "mirror;" our response is an adjustment of what is reflected back to us.[286]

That may mean responding to the targets of predictive policing with investments of resources, rather than the deployment of law enforcement, in the places where those tools forecast crime will take place. It could require that the people we view as potential victims or perpetrators of crimes are treated through a public health lens, rather than a criminal legal system lens, such that we provide those people with an array of services, supports, and investments to ensure that forecasts about them do not come to pass. We may choose to send social workers, doctors, and mental health professionals to respond to forecasts of potential future criminal activity rather than police.[287] Or we may decide to make a different set of investments in those communities expressly focused on supporting institutions that help steer people away from the criminal legal system, such as education, employment, housing, and health.[288]

In the arena of pretrial decision-making, it could be that we calibrate decisions to suggest release for the overwhelming majority of those charged with crimes such that no disparity exists, even if the accuracy of forecasts produced by the tool suffers.[289] We could weigh the outputs used to guide sentencing decisions with data that reflects the nature of racial disparities in

286. *See* Mayson, *supra* note 24, at 2251 (describing the enterprise of prediction through algorithmic tools as a mirror).

287. *See, e.g.*, Christie Thompson, *This City Stopped Sending Police to Every 911 Call*, THE MARSHALL PROJECT (July 24, 2020), https://www.themarshallproject.org/2020/07/24/crisisresponders; Rowan Moore Gerety, *An Alternative to Police that Police Can Get Behind*, THE ATLANTIC (Dec. 28, 2020), https://www.theatlantic.com/politics/archive/2020/12/cahoots-program-may-reduce-likelihood-of-police-violence/617477/; Andy Corbley, *Instead of Responding with Cops, Denver Sends Health Care Teams to Non-Criminal Calls – And it's Already Saving Lives*, GOOD NEWS NETWORK (Feb. 15, 2021), https://www.goodnewsnetwork.org/denver-looks-at-nonviolent-mental-health-policing-with-their-star-social-worker-unit/.

288. *See, e.g.*, Jacyln Cosgrove, *L.A. County Voters Approve Measure J, Providing New Funding for Social Services*, L.A. TIMES (Nov. 3, 2020), https://www.latimes.com/california/story/2020-11-03/2020-la-election-tracking-measure-j (60% of Los Angeles voters voted in favor of Measure J, which "requires that 10% of locally generated, unrestricted county money—estimated between $360 million and $900 million—be spent on a variety of social services, including housing, mental health treatment and investments in communities disproportionally harmed by racism").

289. *See* Yang, *supra* note 124 (explaining that English common law presumed release for those accused of noncapital crimes).

sentencing for particular crimes, communities, and individuals, along with data that accounts for the challenges that individuals face when attempting to reintegrate into society by finding stable employment, housing, healthcare, and other services.

While it could invite a constitutional challenge,[290] we may choose to be race conscious in our responses to the data and tools as we seek to eliminate the racial disparities in the forecasts they produce. Imagine our response to tools that tell us people of color need to be policed more heavily, are more likely to fail to appear in court, or are more likely to recidivate than their white counterparts. We could decide to reduce or eliminate police presence in communities of color, set lower thresholds for pretrial release (or higher bars for pretrial detention), or act in less punitive ways toward all people at sentencing to advance equity. In this way, our response to what the data and the tools are telling us would differ dramatically. System actors employing a race conscious lens could forgo our typical, harsh and punitive responses—too often fueled by race-based inequity in service of a status quo that has always been unfavorable to people of color.

D. Additional Paths Forward

The next Section grapples with three potential policy prescriptions to address the problems presented by algorithmic tools using a racial justice lens. First, such a lens suggests putting the onus on algorithmic tool vendors and system actors to root out and remedy discriminatory impacts imposed by algorithmic tools. Second, it means placing algorithmic tools in the hands of communities to hold accountable those actors who engage in discriminatory or otherwise harmful conduct. Finally, it requires rejecting the type of profiling that actuarial tools encourage. In its place, a racial justice lens suggests adopting an individualized notion of justice that truly accounts for the complexity and story of the person standing before the court, rather than the characteristics that person shared with others.

These solutions seek to shift power to those who are currently powerless given their relationship to the criminal system, while imposing the burdens of antiracism where they belong: on institutional actors and tool vendors. They also raise questions that suggest a broader vision of justice. The hope is that such a dynamic may encourage the type of wholesale transformation the criminal legal system desperately needs, driven by abolition.

---

290. *See infra* notes 397–399 and accompanying text.

1. Shift the Burden

One of the more promising features of algorithmic tools is their ability to surface shortcomings in the law that require a shift in our current legal regime. One such shortcoming is the challenge the law presents in remedying the racially discriminatory harms that individuals may suffer when algorithmic tools are at play. Generally, the onus is on the victim of racial discrimination to use the law in order to identify and remedy their own harm.[291] Unfortunately, the law is not always up to the task. Indeed, the law's failure actually perpetuates the status quo, necessitating a radical intervention to produce progressive change.[292] Placing the burden to root out and remedy algorithmic racial discrimination on tool vendors and the institutions that seek to use them, rather than on those who are assessed by the tools, may be one way to address the law's failure. A review of the constitutional barriers to accountability faced by potential victims of algorithmic discrimination underscores the value of this potential solution.

The Constitution's Equal Protection Clause is the most significant avenue available to challenge the racial discrimination in the administration of criminal justice by state actors.[293] Yet the limits placed on the Equal Protection Clause to redress systemic discrimination in the criminal system have stifled reform and perpetuated inequity for over three decades, since the Supreme Court's decision in *McCleskey v. Kemp*.[294] *McCleskey* applied the purposeful discrimination standard first articulated in *Washington v. Davis*[295] and affirmed in *Village of Arlington Heights v. Metropolitan Housing*

---

291. *See, e.g., McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 802 (1973) (requiring that a complainant charging racial employment discrimination under Title VII prove a prima facie case of racial discrimination. Under this regime, a plaintiff must prove that her race was a "but-for" cause of her adverse treatment).

292. *See* Pat K. Chew & Robert E. Kelley, *Unwrapping Racial Harassment Law*, 27 BERKELEY J. EMP. & LAB. L. 49, 85 (2006) (finding that in racial harassment claims, Black and Asian American plaintiffs have the lowest percentage of wins – at 19.3% and 18.9%, respectively – compared to white plaintiffs, who have a 35% success rate).

293. The focus of this section is on the Fourteenth Amendment's Equal Protection Clause as a means to address discriminatory harms because it is the principal means to do so in the criminal legal system absent state antidiscrimination law. *Whren v. United States*, 517 U.S. 806, 813 (1996) ("[T]he constitutional basis for objecting to intentionally discriminatory application of laws is the Equal Protection Clause. . . ."). Litigants have also unsuccessfully pursued challenges made pursuant to the due process clause. State v. Loomis, 881 N.W.2d 749 (Wis. 2016), *cert. denied sub nom.*, Loomis v. Wisconsin, 137 S.Ct. 2290 (2017).

294. 481 U.S. 279 (1987); Annika Neklason, *The 'Death Penalty's* Dred Scott*' Lives On*, ATLANTIC (June 14, 2019), https://www.theatlantic.com/politics/archive/2019/06/legacy-mccleskey-v-kemp/591424/.

295. 426 U.S. 229, 248 (1976).

Development Corporation.[296] In doing so, the Court ruled that a successful Equal Protection challenge must demonstrate that government officials exercised their discretion with a "discriminatory purpose," amounting to intentional, purposeful discrimination.[297] In the absence of such a determination, an Equal Protection claim must fail.[298] The Court explained that it "would demand exceptionally clear proof before [it] would infer that the discretion has been abused."[299] While the dissent decried the majority's decision as the manifestation of "a fear of too much justice," the intentional discrimination standard has remained the law since 1987.[300]

Equal Protection doctrine is "woeful[ly] inadequa[te]" to address the "forms and dynamics of algorithmic criminal justice tools."[301] First, as a technical matter, it is incredibly difficult to surface intentional discrimination in the context of the tools themselves.[302] "There is no such thing as code that bespeaks racial animus."[303] Evidence of discriminatory intent, difficult to amass when algorithmic tools are not at play, is even more difficult to uncover when trying to assess why particular features of data were selected to train an algorithm.[304] This concern is just one of a number raised by the search of purposeful discrimination in algorithmic tools.[305]

Practical problems of proof also assume that there is explicit malicious intent to be found. It may not be, given that those who design tools and implement them do so with the express intention of addressing the bias that

296. 429 U.S. 252, 270 (1977). While an expansive discussion of the development of Equal Protection jurisprudence is beyond the scope of this article, a brief review is useful. The Supreme Court's jurisprudence shifted in the late 1970s, from examining allegations of purposeful discrimination in context to searching for malice. The Court turned away from a concern with purposeful discrimination through the lens of "contextual intent," which "focused on motives [of alleged discriminatory actors] only in the loosest sense (and sometimes not at all)." Ian Haney-López, *Intentional Blindness*, 87 N.Y.U. L. REV. 1779, 1785 (2012). Abandoning context, in subsequent years, the Court turned to a search for "malicious intent" which "declares direct proof of injurious motives a prerequisite and, more pertinently, renders contextual evidence irrelevant." *Id.*

297. *McCleskey*, 481 U.S. at 298.

298. *Id.*

299. *Id.* at 297.

300. *Id.* at 339 (Brennan, J., dissenting).

301. Huq, *supra* note 11, at 1083.

302. *Id.* at 1102.

303. *Id.* at 1066.

304. *Id.* at 1098.

305. Others include sorting out and assigning malice to the motives for the design and implementation of an algorithmic tool, looking at the challenges posed by examining the wide range of actors in the system that contribute to the data that informs the tools and "aggregating a large number of dispersed individual motives so as to ascertain whether a but-for standard of intentionality has been met by a collectivity," and considering whether reliance on flawed data would amount to intentional discrimination. *Id.* at 1088–94.

so often pervades the criminal legal system, even when those tools encourage decisions that may reflect the biased data that they are fed.[306] Unlike the police officer who explicitly engages in racial profiling, or the prosecutor who exercises discriminatory peremptory challenges, or the judge who, with purposeful animus, levies harsh punishments on people of color, the motives of those in the algorithmic tool business are publicly stated as racially benevolent.[307] That benevolence, in the context of a legal framework designed to respond to explicit and malicious acts of racial discrimination, is a shield from the interrogation that proof of an Equal Protection violation requires.[308]

In many ways, the concern that the searching scrutiny of our Equal Protection framework fails to account for the way actuarial tools operate parallels concerns first raised by Professor Charles Lawrence in his seminal work highlighting the gap between "unconscious bias" and the purposeful discrimination standard imposed by the Supreme Court's interpretation of the Equal Protection Clause.[309] The advent of algorithmic tools and the regulation that accompanies their use provides a new opportunity to upset old standards that have proven unresponsive to the realities of discrimination.[310]

---

306. *See Risk Assessments, When Paired with Appropriate Policies, Can Contribute Significantly to Pretrial Reform*, ARNOLD VENTURES (July 1, 2019), https://www.arnoldventures.org/newsroom/risk-assessments-when-paired-with-appropriate-policies-can-contribute-significantly-to-pretrial-reform/ ("We are strongly committed to reducing racial bias in pretrial decision making. In particular, we seek to understand how risk assessment can be used to reduce racially disparate outcomes.").

307. *Id.*; *see also* Ferguson, *supra* note 115 (describing a predictive policing company's efforts to account for racial bias in policing).

308. Huq, *supra* note 11, at 1088 ("The concerns of constitutional law simply do not map onto the ways in which race impinges on algorithmic criminal justice. The result is a gap between legal criteria and their objects. Crucially, the two main doctrinal touchstones of bad intent and bad classifications provide scant traction for the analysis of algorithmic criminal justice. Both hinge on concepts that translate poorly, if at all, to the algorithmic context and are not easily adapted for application to that end. A focus on racial animus will almost never be fruitful. A focus on classification leads to perverse and unjustified results. The replacement of unstructured discretion with algorithmic precision, therefore, thoroughly destabilizes how equal protection doctrine works on the ground. The resulting mismatches compel my conclusion that a new framework is needed for thinking about the pertinent racial equity questions.").

309. *See* Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317, 318–26 (1987).

310. Even efforts to hold accountable those who use actuarial tools with a knowledge of their disparate impact are foreclosed by the law, because "'[d]iscriminatory purpose' . . . implies more than intent as volition or intent as awareness of consequences. It 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." McCleskey v. Kemp, 481 U.S. 279, 298 (1987) (quoting Personnel Adm'r of Massachusetts v. Feeney, 442 U.S. 256, 279 (1979)) .

Given this opportunity to change the standard, policymakers must act. They must craft new regulatory schemes that can vindicate the potential harms imposed by algorithmic tools and fill the gaps of the Equal Protection Clause. At a minimum, such legislation should impose an ex ante check on algorithmic tools to alleviate harmful disparate impacts and to ensure that there is continued monitoring of the potentially harmful burdens imposed by use of the tools.[311] Such a framework has the potential to alleviate the challenges of proof that litigants face in demonstrating that the harmful effects of a tool go beyond an individual to others who are similarly situated.

A cursory survey of legislative activity and advocacy efforts attempting to curb the discriminatory harms imposed by actuarial justice provides some encouragement. New York has enacted a requirement that pretrial risk assessments be "designed and implemented in a way that ensures the results are free from discrimination on the basis of race, national origin, sex, or any other protected class."[312] Notably, while this provision has not yet been applied or interpreted by any New York courts, it contains no explicit intent requirement. It also imposes an affirmative obligation on the state to ensure that the tools they use are free from discrimination.[313] Similar legislation was under consideration in Washington and enacted into law in Idaho.[314] At the federal level, three members of Congress introduced the Algorithmic Accountability Act of 2019, which would essentially require technology vendors to test the algorithms they use for bias.[315] Once again, this legislation

311. Scholars have offered ways to measure the impact of algorithmic tools on racial equity. Huq, *supra* note 11, at 1128 ("[A]n appropriate benchmark would home in upon the net cost (or benefit) of an algorithmic criminal justice instrument for the racial minority in the socially subordinate position."). My concern is not so much with the metric of fairness being used, though that is deeply important, but requiring those who seek to design and implement the tools to demonstrate that they do not exacerbate racial inequality. Choice of a fairness metric is a policy determination, rather than a technical one, that would need to be made in the policymaking process. *See* Mayson, *supra* note 24 at 2238–47 (detailing a host of applicable equality metrics); *see id*. at 2294–95 (describing a combination of equality metrics that tools might meet); *see also* THE LEADERSHIP CONFERENCE ON CIVIL AND HUMAN RIGHTS, THE USE OF PRETRIAL "RISK ASSESSMENT" INSTRUMENTS: A SHARED STATEMENT OF CIVIL RIGHTS CONCERNS (2018), http://civilrightsdocs.info/pdf/criminal-justice/Pretrial-Risk-Assessment-Full.pdf (recommending the use of varied measures of racial equity).

312. N.Y. CRIM. PROC. LAW § 510.45(3)(b)(i) (McKinney 2020).

313. *Id.* § 510.45(3)(b)(i).

314. DJ Pangburn, *Washington Could Be the First State to Rein in Automated Decision-Making*, FAST CO. (Feb. 18, 2019), https://www.fastcompany.com/90302465/washington-introduces-landmark-algorithmic-accountability-laws; Beryl Lipton, *Idaho Legislators Approve Law Requiring Transparency for Risk Assessment Tools*, MUCKROCK (Mar. 26, 2019), https://www.muckrock.com/news/archives/2019/mar/26/algorithms-idaho-bill-update/.

315. Algorithmic Accountability Act of 2019, H.R. 2231, 116th Cong. (2019). There are, in fact, efforts underway to audit algorithmic tools in a range of domains, though the parameters of the audit, the undefined nature of the field, the ways in which private companies choose to deploy the

imposes an affirmative obligation on those who seek to design and implement tools.

Policy advocates have also advanced frameworks to shift the burden of rooting out harm from individuals to stakeholders. Algorithmic Impact Assessments ("AIA") are one such example of this burden shifting framework. Modeled on environmental impact assessments, AIAs work by requiring government agencies to "assess how . . . systems are used, whether they are producing disparate impacts, and how to hold them accountable."[316] They require government agencies to conduct a self-assessment of existing and proposed algorithmic tools to evaluate their potential impacts, engage external researchers to conduct ongoing auditing, publicly disclose audit results prior to procurement of an algorithmic tool, solicit public comments regarding the tool, and provide mechanisms for communities or individuals to challenge systems that produce harms.[317] The framework is meant to enhance public accountability of algorithmic tools, "[i]ncrease public agencies' internal expertise and capacity to evaluate the systems they build or procure" for disparate impacts, and empower the public with knowledge about tools in use and opportunities to determine the contours of accountability.[318]

Assessments of impact prior to adoption and implementation are already required in some jurisdictions with regard to criminal justice policy. Racial impact statements, for example, allow lawmakers to evaluate the racial disparities that legislation may produce before it is adopted and implemented.[319] In 2008, Iowa became the first state to adopt such a measure,

---

audit, and the relatively small number of firms that offer auditing capabilities provide some sense of the challenges posed by those efforts. Alfred Ng, *Can Auditing Eliminate Bias from Algorithms?*, MARKUP (Feb. 23, 2021), https://themarkup.org/ask-the-markup/2021/02/23/can-auditing-eliminate-bias-from-algorithms.

316. Dillon Reisman et. al., *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, AINOW 1, 4 (2018), https://ainowinstitute.org/aiareport2018.pdf.

317. *Id.*

318. *Id.* at 5.

319. Nicole D. Porter, *Racial Impact Statements*, SENT'G PROJECT (Sept. 30, 2019), https://www.sentencingproject.org/publications/racial-impact-statements/. In the algorithmic context these efforts could take on the character of race audits, described by Professor Robin Lenhardt as evaluative mechanisms that can be used "by localities interested in grappling with the inequalities that attend the color line." R.A. Lenhardt, *Race Audits*, 62 HASTINGS L.J. 1527, 1534 (2011). Such audits do not search for "the proverbial wrongdoer," but instead surface how racial inequality reveals itself in systems, procedures, practices, and relationships of a municipality across multiple life domains. *Id.* Subjecting algorithmic tools to a similar audit, sensitive to racial inequality, would theoretically expose how an algorithmic tool might perpetuate inequality. *See also,* Deborah N. Archer, *"White Men's Roads Through Black Men's Homes": Advancing Racial Equity Through Highway Reconstruction*, 73 Van. L. Rev. 1259, 1321 (2020) (recommending the use of racial equity impact studies by "policymakers embarking on highway development and

with four other states doing so since then.[320] The Minnesota Sentencing Guidelines Commission produces racial impact statements, though is not required to do so by law.[321] New Jersey became the latest state to do so in 2018, passing a law that "requires the state's Office of Legislative Services to prepare racial-impact statements for policy changes that affect pretrial detention, sentencing and parole."[322] Such measures allow jurisdictions to uncover the causes of racial disparities and to understand how policy changes can exacerbate or reduce them.[323]

Other accountability and oversight measures may rely more heavily on vendors and private industry. Ethical codes of conduct that impose moral commitments on those who produce technology are another means of oversight, though they may place too much reliance on the malleable moral compass of corporate actors, rendering such measures unreliable.[324] Requiring an algorithm's proponent to provide a human impact statement, which could outline the expected ramifications of an algorithmic tool on a population, is one other accountability mechanism.[325]

In practice, such measures might require algorithmic tool designers to disclose the datasets they relied upon to develop their tools, the efforts they undertook to assess those datasets for biases, and the measures taken to ensure that the forecasts produced by the tools do not unjustifiably vary by race.[326] Such requirements are perfectly reasonable and well understood in

---

redevelopment projects should engage in a systematic, comprehensive, and holistic review of how racial and ethnic groups will be impacted by the project")

320. Porter, *supra* note 319.

321. *Id*.

322. *Id*.

323. Barkow, *supra* note 10, at 1610–13.

324. Sonia Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 108 (2019) ("The issue of algorithmic accountability demonstrates one core aspect that is missing among computer scientists and software engineers: a concrete, user-friendly, ethical platform with which to approach decisionmaking and software design.").

325. *See id*. at 115–18 (describing the proposed elements of a human impact statement); *see also* Erin Murphy, *The Mismatch Between Twenty-First-Century Forensic Evidence and Our Antiquated Criminal Justice System*, 87 S. CAL. L. REV. 633, 658–61 (2014) (recommending a "collective confrontation right to transparency and accountability standards in forensic analysis" that would place the onus on proponents of forensic evidence to offer evidence of structural and systemic features and quality control measures that ensure the accuracy of such evidence).

326. This work would require grappling with, and potentially using, different measures of fairness to evaluate unwarranted disparities produced by an algorithmic tool. See *supra* notes 244–256 and accompanying text for a discussion of two measures of fairness. It would also require vendors to examine the data used, the algorithm, and the outputs for racial disparities. One can imagine a range of efforts that vendors may have to undertake to account for unwarranted racial disparities, including consulting additional data sources, disregarding data sources, weighting forecasts produced by their tools, or providing an explicit disclaimer about the reliability of the output forecast because racial inequality is baked into the data relied upon. Forcing proponents of

administrative law.[327] Fundamentally, these frameworks force those who seek to develop and wield algorithmic tools to ask the difficult questions about the racialized harms they may produce upfront and actually address those harms when they surface. They must also accommodate the critiques leveled at tools by impacted communities. In their absence, those at the greatest risk of suffering a racially disparate impact are left to the inadequate tools provided by the law, ensuring that bias will persist, largely unchecked.

2. Flip the Gaze

Acknowledgment of racism's permanence helps to shape the contours of another response to the concerns driven by actuarial and algorithmic decision-making: turning the tools away from the individuals subjected to the system and toward the institutional actors who run it. The idea behind this recommendation is simple, and in part stems from the notion that interrogating the role of race and racism in inequality is essential. If institutional actors and reform advocates really want to address unwarranted disparities in the administration of justice, we must be willing to subject those whose decisions shape the system to the same data-driven, evidence-based scrutiny that we foist on the people being shuffled through it. What is good for the goose is good for the gander.

This inversion of the target carries with it several potential benefits. Institutional actors—in particular, those who are making judgments about individuals—use algorithmic tools in the hopes that it will allow them to properly sort and classify individuals and make better decisions about them. Why not apply that same logic to the decisionmakers themselves? Understanding their behavior requires that we track the decisions they make when faced with a certain set of facts, particular pieces of information, or specific types of people. That understanding can foster accountability through transparency by exposing the points where emotion, unreasonable risk aversion, or flawed judgments override facts and evidence to the detriment of those being judged.[328]

---

algorithmic tools to explain themselves carries with it the potential benefits of improving the quality of outcomes and "deter[ring] bias and arbitrariness." Katherine J. Strandburg, *Rulemaking and Inscrutable Automated Decision Tools*, 119 COLUM. L. REV. 1851, 1868 (2019). There are a host of decision points that could be interrogated and disclosed for these purposes. *See id.* at 1872–73 (describing "aspects of the development of machine-learning-based decision tools, and of the decision rules embedded in those tools, that are . . . explainable").

327. Strandburg, *supra* note 326, at 1882–84 (crafting a framework to promote transparency of machine learning tools by interpreting and applying administrative law practices to require an explanation of a tool and preservation of information about the source of the training data, its selection, and the methods used to validate the tool).

328. *See* Sarah Brayne, PREDICT AND SURVEIL: DATA, DISCRETION, AND THE FUTURE OF POLICING 101–06 (2020) (explaining how data can be used to reduce inequality in the criminal legal

It might also engender some empathy on the part of those institutional actors who reflexively assert that subjecting people who are entangled with the criminal legal system to actuarial decision-making is the best path forward.[329] Experience can be sobering. No one really likes to be held to account, monitored, or have their decisions called into question.[330] First, there is the general uneasiness that comes with being surveilled, tracked, and having one's privacy upended.[331] Compounding that is the fear that one's hard-earned, experience-driven, professional judgment is devalued by the introduction of algorithmic tools.[332] Changing the targets might lead to more creative thinking about the efficacy of algorithmic tools and the value of their forecasts.

Finally, it allows us to use data to shape the discretion exercised by system actors. Commendable behavior can be encouraged by providing support and guidance as needed where discriminatory or otherwise harmful decision-making has surfaced. Those supports could include changes to policy or practice, training on bias and decision-making, or a narrowing of the choices available to avoid poor decision-making. These remedial efforts need not be punitive if there is broad-based commitment to cleansing the criminal legal system of as much injustice and unfairness as possible.[333] Although a simple agreement about the need for change may not be enough to overcome the natural resistance put forward by system actors, that

---

system by "aggregating data on police practices [to] shed light on systematic patterns and institutional practices previously dismissed as individual-level bias, ultimately providing an opportunity to police the police. . . .").

329. While one would hope decision makers subjected to algorithmic tools would empathize with those who face such tools in the criminal legal system, there is no guarantee that they would. *See id*. at 98 (detailing the cognitive dissonance of police officers who did not recognize the parallels between their discomfort with managerial surveillance and the surveillance technologies they imposed on others).

330. Brayne & Christin, *supra* note 169, at 4, 9; *see* Brayne, *supra* note 328 at 75–99 (describing law enforcement reaction to, and resistance against, managerial surveillance).

331. Brayne & Christin, *supra* note 169, at 9; *see also* Harry Bruinius, *Why Police Are Pushing Back on Body Cameras*, CHRISTIAN SCI. MONITOR (Aug. 30, 2016), https://www.csmonitor.com/USA/Justice/2016/0830/Why-police-are-pushing-back-on-body-cameras.

332. Brayne & Christin, *supra* note 169, at 9.

333. This attitude is reflected in civil and human rights lawyer Bryan Stevenson's reminder that part of the challenge of addressing racial inequality is that American society is obsessed with punishment. That obsession has led people to feel that by surfacing the ways that racial inequality has infected the world around us, punishment must result. As Stevenson reminds us,

> I'm not interested in prioritizing punishment. I want to liberate us. I want to get to the point where we can say, "'That was bad and that was wrong and we need to get to someplace that's better!"' I want to deal with this smog created by our history of racial inequality, so we can all breath something healthy, feel something healthy.

THOMPSON ET. AL., *supra* note 268, 88.

resistance should not upend efforts to impose transparency and accountability.

Two conditions of flipping the gaze of algorithmic tools are essential. First, the effort to close the racial equity gap by using algorithmic tools must be informed by a harm reduction framework, rather than the type of ratcheting up that results when the answer to inequity is to treat everyone more harshly. For example, if a tool reveals that a judge imposes more lenient sentences on white people than Black people, resolving that disparity would require treating all people with the same type of leniency that white people receive, rather than sentencing more white people to lengthier terms of incarceration. Ultimately, imposing such a condition helps to avoid the same dynamic that often taints criminal policymaking to "reward punitiveness and punish mercy."[334]

Second, the tools must be placed in the hands of communities empowered to hold institutions and actors accountable. "Community" is defined here as the network of individuals who are advocating for equity in the criminal legal system and are bound together by their common concerns about the inequities fostered by the use of algorithmic tools and the criminal legal system. Failing to provide community control has the potential to wholly undermine the value of flipping the gaze.[335]

The experience of body-worn cameras as accountability mechanisms for police conduct provides a ready example of what happens when the hands

334. Maurice Chammah, *Could Removing Brock Turner's Judge Hurt Poor and Minority Defendants?*, MARSHALL PROJECT (June 16, 2016), https://www.themarshallproject.org/2016/06/16/could-removing-brock-turner-s-judge-hurt-poor-and-minority-defendants (describing how the removal of a judge following public backlash against judge's perceived leniency in sentencing a white youth in a sexual assault case may discourage leniency and lead to more severe sentences for clients of color convicted of similar offense); *see also* Barkow, *supra* note 3, at 105–23 (describing how the public, elected officials, and interest groups advance tough on crime policies that favor lengthy sentences).

335. Community education is another critical component of community oversight of system actors via technology. It is essential that communities understand the relationship between technology, the data upon which it relies, and the ways that systems and institutions function. A prominent example of this educational work is being done by the Our Data Bodies Project, which describes itself as "a five-person team concerned about the ways our communities' digital information is collected, stored, and shared by government and corporations." *Who We Are*, OUR DATA BODIES: HUM. RTS. AND DATA JUS., https://www.odbproject.org/about/who-we-are/ (last visited Jan. 15, 2020). Given that concern, the Project focuses on the intersection of data collection and human rights, and provides guidance on data protection, supports community education and organizing, and demonstrates how data impacts domains such as housing, urban development, public benefits, and reentry. *Id.* In a somewhat similar vein, the Detroit Digital Justice Coalition hosts workshops called DiscoTechs, short for Discovery Technology. *About*, DETROIT DIGIT. JUST. COAL., http://detroitdjc.org/about/story/ (last visited Jan. 29, 2021). These workshops "are a space to learn about the impact and possibilities of technology within our communities," and serve to "demystify, engage, and inform the community about issues of Internet use and ownership, and our communications rights on and offline." *Id.*

that wield the tools are unchanged.[336] Body-worn cameras were widely adopted to curtail police violence against communities of color.[337] Isolated incidents of success meant that they quickly were adopted as part of the standard suite of remedial mechanisms in systemic efforts to reform policing.[338] Lawsuits and consent decrees demanded their use.[339] Police departments nationwide, spurred on by the promise of additional federal funding, acquired them.[340]

Yet for all of the accountability promised, the institutional actors holding the tools of accountability have not changed, which means the tools have not been able to meet their potential.[341] We have not seen a wholesale

336. Amanda Ripley & Timothy Williams, *Body Cameras Have Little Impact on Police Behavior, Study Says*, N.Y. TIMES (Oct. 20, 2017), https://www.nytimes.com/2017/10/20/us/police-body-camera-study.html; German Lopez, *The Failure of Police Body Cameras*, VOX (July 21, 2017), https://www.vox.com/policy-and-politics/2017/7/21/15983842/police-body-cameras-failures.

337. Lindsey Van Ness, *Body Cameras May Not Be the Easy Answer Everyone Was Looking For*, PEW (Jan. 14, 2020), https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2020/01/14/body-cameras-may-not-be-the-easy-answer-everyone-was-looking-for (describing the widespread adoption of body-worn cameras following the 2014 killing of Michael Brown at the hands of a police officer); Ben Miller, *Data Pinpoints the Moment When Police Body Cameras Took Off*, GOV'T TECH. (Jan. 28, 2019), https://www.govtech.com/data/Data-Pinpoints-the-Moment-When-Police-Body-Cameras-Took-Off.html.

338. A 2016 Bureau of Justice Statistics survey concluded that nearly half of U.S. law enforcement agencies had acquired body-worn cameras. Shelley S. Hyland, BUREAU OF JUSTICE STATISTICS, BODY-WORN CAMERAS IN LAW ENFORCEMENT AGENCIES (2018), https://www.bjs.gov/content/pub/pdf/bwclea16.pdf; Cynthia Lum et. al., *Research on Body-Worn Cameras: What We Know, What We Need to Know*, CRIMINOLOGY & PUB. POL'Y 93, 94 (2019), https://onlinelibrary.wiley.com/doi/epdf/10.1111/1745-9133.12412 ("Body-worn cameras . . . are one of the most rapidly diffusing technologies in policing today. . . .").

339. New York's experience with body-worn cameras is one example. Ashley Southall, *New York's First Police Body Cameras Take to Streets in Upper Manhattan*, N.Y. TIMES (Apr. 27, 2017), https://www.nytimes.com/2017/04/27/nyregion/new-york-police-department-body-cameras.html (explaining that introduction of body-worn cameras was part of the remedies set forth in litigation regarding the unconstitutional policing tactics of the New York City Police Department); *Floyd v. City of New York*, 959 F. Supp. 2d 668, 685 (S.D.N.Y. 2013) ("Because body-worn cameras are uniquely suited to addressing the constitutional harms at issue in this case, I am ordering the NYPD to institute a pilot project in which body-worn cameras will be worn for a one-year period by officers on patrol in one precinct per borough—specifically the precinct with the highest number of stops during 2012.").

340. *See supra* notes 337–339.

341. *Body Worn Camera Basics*, BALT. POLICE DEP'T (last visited May 7, 2021), https://www.baltimorepolice.org/transparency/body-worn-cameras (announcing that of 133,000 videos recorded in first six months of implementation of body worn cameras across the Baltimore Police Department, fourty-seven were flagged for review of potential misconduct); Megan Hickey, *How Often Do Chicago Police officers Fail to Activate Their Body Cameras? It's Hard to Know*, CBS 2 CHI. (July 30, 2019) https://chicago.cbslocal.com/2019/07/30/inspector-general-chicago-police-body-cameras/ (reporting that an investigation by the City of Chicago Office of the Inspector General found that lieutenants failed to review body camera footage or discipline officers who did

change in the culture of policing or definitive evidence that the cameras reduce police use of force.[342] That can be explained by who holds the tools and power to ensure accountability.[343] It is often the police who decide when to operate the cameras and what gets recorded.[344] Before an incident reaches a prosecutor's desk, a myriad of hurdles may stand in the way of accountability. As a policy matter, the cameras, the data, and the footage they produce are securely in the possession of law enforcement until law enforcement decides to make it public.[345] And police, along with other system actors, may have an outsized say over whether conduct caught on camera will warrant a corrective intervention. Prosecutors rarely prosecute police.[346] That fact does not change when prosecutors do get body camera footage, which is far more often used to prosecute civilians.[347] Thus, rather than providing a community with an accountability measure, the cameras have served as another point of grievance by the community seeking accountability.[348]

---

not comply with the body worn camera policy, thereby violating the federal consent decree that mandated and funded the department's body worn camera policy).

342. Lum et al., *supra* note 338, at 109

343. *See* Amna A. Akbar, *Toward a Radical Imagination of Law*, 93 N.Y.U. L. REV. 405, 465–66 (2018) (describing the cameras as "technology that remains in the hands of the police and at the mercy of the prosecutor [and] remains embedded in a criminal system bureaucracy that has more interest in protecting itself than in accountability for its violence against Black people").

344. *See Police Body Camera Policies: Recording Circumstances*, BRENNAN CTR. FOR JUST. 1 (Aug. 3, 2016), https://www.brennancenter.org/our-work/research-reports/police-body-camera-policies-recording-circumstances

345. Chad Marlow & Gary Daniels, *Ohio Bucks a Bad Trend With New Police Body Camera Law*, ACLU (Feb. 5, 2019, 10:15 AM), https://www.aclu.org/blog/privacy-technology/surveillance-technologies/ohio-bucks-bad-trend-new-police-body-camera-law/.

346. *Id. But see* Justin Fenton, *Baltimore Police Officer Who Turned Off Body Camera Charged With Tampering With Evidence; Others Cleared*, BALT. SUN (Jan. 24, 2018), https://www.baltimoresun.com/news/crime/bs-md-ci-body-camera-tampering-20180124-story.html.

347. One study found that 93% of prosecutors who responded that their jurisdiction uses body worn cameras use the footage from those cameras primarily to prosecute citizens rather than police. Lum et al., *supra* note 338 at 108.

348. For thirteen months, Chicago Mayor Rahm Emmanuel blocked the release of dashboard camera footage showing Chicago police officer Jason Van Dyke killing Laquan McDonald, a Black seventeen-year-old. While Van Dyke claimed that McDonald lunged at him with a knife, the video shows Van Dyke shooting McDonald sixteen times as McDonald walked away from him. The delay in releasing the video coincided with Emmanuel's reelection campaign. Jessica Glenza, *Chicago Officials Delayed Release of Laquan McDonald Shooting Video*, GUARDIAN (Jan. 1, 2016) https://www.theguardian.com/us-news/2016/jan/01/chicago-officials-delayed-release-laquan-mcdonald-shooting-video; Bernard E. Harcourt, *A Cover-Up in Chicago*, N.Y. TIMES (Nov. 30, 2015) https://www.nytimes.com/2015/11/30/opinion/cover-up-in-chicago.html. Van Dyke was eventually charged with murder, convicted, and sentenced to nearly seven years in prison. Mitch Smith & Julie Bosman, *Jason Van Dyke Sentenced to Nearly 7 Years for Murdering Laquan McDonald*, N.Y. TIMES (Jan. 18, 2019), https://www.nytimes.com/2019/01/18/us/jason-van-dyke-

That is why community control and power are so important. When the gaze of technology is flipped on system actors and the tools are placed in community hands, it can be used to interrogate and evaluate the system in ways that are innovative and beneficial.[349] For example, Campaign Zero—an organization dedicated to reducing police violence nationwide[350]—began using big data to evaluate California's 100 largest municipal police departments based on the number of arrests made for low-level offenses, the use of force during an arrest, the rate of homicides solved, the presence or absence of racial disparities in arrests and use of force, and the treatment of civilian complaints of police abuse.[351] That information can be used to advocate for changes to police policies and practices.

Chicago's Citizens Police Data Project also provides a measure of accountability through data by "tak[ing] records of police interactions with the public—records that would otherwise be buried in internal databases—and opens them up to make the data useful to the public, creating a permanent record for every . . . police officer."[352] CAPstat, a police accountability database modeled on the Chicago tool, has been developed and is in use in New York.[353] Relying on publicly available data collected from various sources between 2011 and 2018, the database demonstrates:

> [T]ransparency can improve our collective ability to identify trends of misconduct across, for example, different types of allegations,

sentencing.html. While Chicago officials have moved more quickly since 2015 to release body camera and dashboard camera footage, the process of releasing such footage varies state by state, department by department. Richard Fausset & Giulia McDonnell Nieto del Rio, *As Body Cameras Become Common, a Debate Over When to Release the Footage*, N.Y. TIMES (May 2, 2021) https://www.nytimes.com/2021/05/02/us/police-body-cameras-andrew-brown-north-carolina.html. For example, although protesters have demanded the release of body camera footage depicting the police killing of Andrew Brown Jr., a judge has delayed its release for at least 30 days. Richard Fausset & Giulia McDonnell Nieto del Rio, *Judge Declines Immediate Release of Video in North Carolina Shooting*, N.Y. TIMES (Apr. 30, 2021) https://www.nytimes.com/live/2021/04/28/us/andrew-brown-jr-shooting-body-camera.

349. *See* Mayson, *supra* note 24, at 2284–86 (recommending the use of algorithmic tools as diagnostic measures to identify sites of racial disparity for correction).

350. *See Vision*, CAMPAIGN ZERO (last visited March 1, 2021), https://www.joincampaignzero.org/#vision.

351. *Police Scorecard*, CAMPAIGN ZERO, https://policescorecard.org/findings (last visited Jan. 15, 2020).

352. *Citizens Police Data Project*, INVISIBLE INST., http://invisible.institute/police-data (last visited Jan. 17, 2020). Critically, it was this data set that revealed multiple allegations of misconduct against Chicago Police Officer Chicago police officer Jason Van Dyke. Sarah Kaplan, *Chicago Police Officer Charged in Deadly Shooting Has a History of Misconduct Complaints*, WASH. POST (Nov. 25, 2015), https://www.washingtonpost.com/news/morning-mix/wp/2015/11/25/chicago-cop-charged-in-deadly-shooting-has-a-history-of-misconduct-complaints/; *see also supra* note 348 and accompanying sources.

353. *What is this Data?*, CAPSTAT, https://www.capstat.nyc/about/what/ (last visited Jan. 17, 2020).

> commands and units that could inform policy debates, improve public discourse about police misconduct allegations and be a resource for people who witnessed or were harmed by police misconduct to help them decide what to do next.[354]

Though these tools specifically targeted police behavior, one can readily imagine a similar tool focused on the conduct of judges, prosecutors, defense attorneys, probation officials, and parole officials. The Vera Institute for Justice demonstrated this concept through a project aimed at addressing prosecutorial discretion in Milwaukee, Wisconsin, and Mecklenburg County, North Carolina.[355] That work involved collecting data to monitor the exercise of discretion by prosecutors at various decision points in the criminal legal system.[356] Researchers then analyzed that data to determine the source of racial disparities for particular charging decisions associated with drug crimes.[357] Once researchers identified the sources of disparity by examining the data and engaging in a qualitative analysis of decision-making,[358] prosecutors in Milwaukee discovered that junior, less experienced prosecutors pursued drug paraphernalia cases more aggressively than their colleagues.[359] Mecklenburg County prosecutors found that drug paraphernalia cases constituted 97% of all drug cases, and "press[ing] [charges] for all drug cases and every drug charge" involving Black women, despite the fact that many of those cases were ultimately dismissed or resolved with a diversion into drug treatment.[360] These offices made policy changes to address their findings, resulting in a narrowing of racial disparities for a subset of the crimes prosecuted by both offices.[361]

Another example can be found in a recent analysis of just over 105,000 criminal cases handled by the Legal Aid Society of New York and the bail decisions made by judges in those cases.[362] Although the data collected did

354. *Id.*

355. Wayne McKenzie et al., VERA INST. OF JUST., PROSECUTION AND RACIAL JUSTICE: USING DATA TO ADVANCE FAIRNESS IN CRIMINAL PROSECUTION 1 (Mar. 2009), https://www.vera.org/downloads/Publications/prosecution-and-racial-justice-using-data-to-advance-fairness-in-criminal-prosecution/legacy_downloads/Using-data-to-advance-fairness-in-criminal-prosecution.pdf.

356. *Id.* at 5.

357. *Id.* at 6–7.

358. *Id.*

359. *Id.*

360. *Id.*

361. *Id*.

362. Anna Maria Barry-Jester, *You've Been Arrested. Will You Get Bail? Can You Pay It? It May All Depend on Your Judge.*, FIVETHIRTYEIGHT (Jun. 19, 2018), https://fivethirtyeight.com/features/youve-been-arrested-will-you-get-bail-can-you-pay-it-it-may-all-depend-on-your-judge/.

not allow for any demographic analysis, what the effort uncovered was that bail amounts set on people accused of crimes varied dramatically based on where a person was arraigned, the crime they were charged with and the judge presiding over the arraignment.[363] Getting arrested on the "wrong day" could mean that a person is "more than twice as likely to have to" post bail to purchase their freedom.[364]

These examples demonstrate how collecting data about the past behavior of system actors can help inform how they might behave going forward in ways that align with racial equity. All that is required is a willingness to subject system actors to scrutiny.[365] That is no easy feat; but if accomplished, greater scrutiny would shift the nature of the inquiry undertaken by algorithmic tools in the criminal legal system.

3. Listen to the People Being Judged.

One feature of algorithmic decision-making is that it emphasizes quantitative data for predictive purposes over the narratives that shape the lives of the individuals to be judged by the state.[366] That emphasis is troubling because it is done in service of what amounts to profiling— though it is often characterized as prediction. Rather than give in to that dynamic, it is necessary to adopt an orientation that views stories and anecdotes as data points that carry just as much—if not more—power than raw numbers and leave predictive analytics behind.[367] The march toward algorithmic tools

363. *Id.*

364. *Id.*

365. *See* Roberts, *supra* note 27, at 1726–28 (recommending that people employ technology to "identify and excavate the sites where inequality has been institutionally embedded").

366. *See supra* Section I.D.3.

367. In discussing the power of stories to "destabilize hardened and assumed norms" one scholar pointed to examples of stories told by Black women enmeshed in the criminal legal system, explaining that:

> personal narratives reveal types of information and knowledge that are neither manifested in the doctrinal representations of their stories nor necessarily reflected in the statistics that present the quantitative picture of black women within the criminal justice system. If nothing else, both the statistics pertaining to the conviction and incarceration rates of African-American women discussed below and stories . . . remind us that there is a real cost to being marked by difference within society. Telling our versions of our stories is merely a first step in revealing the reach of institutional power and the systemic nature of oppression.

Mario L. Barnes, *Black Women's Stories and the Criminal Law: Restating the Power of Narrative,* 39 U.C. DAVIS L. REV. 941, 954, 957 (2006). Indeed, stories open up the world in ways that can alter human judgment:

> Stories humanize us. They emphasize our differences in ways that can ultimately bring us closer together. They allow us to see how the world looks from behind someone else's spectacles. They challenge us to wipe off our own lenses and ask, "Could I have been overlooking something all along?" Telling stories invests text with feeling, gives voice

needs to be balanced by an accounting of the context that produces the quantitative data.

A racial justice lens suggests an approach that privileges the voices and narratives of those closest to the harms perpetuated by the system.[368] That does not mean dispensing with numbers altogether. Statistics serve real and important purposes. They can inform decision-making, shape policy, and highlight patterns of harm. They can help decisionmakers take stock of the barriers that stand in the way of the accused in a more rigorous way. They can also foster transparency around the decision-making process and provide an avenue for accountability. Indeed, the entire premise of flipping the gaze of tools on the actors in the system is rooted in the idea that statistics bear value and can influence the exercise of discretion.

These benefits carry a danger. Privileging quantitative data over qualitative information can blind decisionmakers from doing the work to uncover the unique forces, facts, and circumstances that lead people into the criminal legal system. Pure reliance on statistical metrics stifles the curiosity and creativity that system actors may need to fully engage the complexities of peoples' lives in meaningful, productive and effective, ways. For example, the number of individuals who fail to appear in court following their initial release from custody is a specific data point. That information can shape systemic responses to failures to appear. But without context and nuance—answering the question of why it is that people fail to appear in court—those responses will be inadequate. Stories—the qualitative information—give meaning to the numbers.

There is another important benefit to stories. One of the many things that I learned during my time defending people who were accused or

> to those who were taught to hide their emotions. Hearing stories invites hearers to participate, challenging their assumptions, jarring their complacency, lifting their spirits, lowering their defenses. Stories are useful tools for the underdog because they invite the listener to suspend judgment, listen for the story's point, and test it against his or her own version of reality.

Richard Delgado, *Storytelling for Oppositionists and Others: A Plea for Narrative*, 87 MICH. L. REV. 2411, 2440 (1989). Accordingly, "[s]tories, parables, chronicles, and narratives are powerful means for destroying mindset — the bundle of presuppositions, received wisdoms, and shared understandings against a background of which legal and political discourse takes place." *Id.* at 2413.

368. "[T]hose who have experienced discrimination speak with a special voice to which we should listen. Looking to the bottom—adopting the perspective of those who have seen and felt the falsity of the liberal promise—can assist critical scholars in the task of fathoming the phenomenology of law and defining the elements of justice." Mari J. Matsuda, *Looking to the Bottom: Critical Legal Studies and Reparations*, 22 HARV. C.R.-C.L. L. REV. 323, 324 (1987). That is because those who have been harmed by the oppressive, interlocking systems of racial power have unique experiences and insights to offer that generate novel solutions to racial inequality. *Id.* at 325.

convicted of crimes is that context matters. I came to learn that when judges or other actors in the criminal legal system took stock of that context, more often than not it worked to the benefit of my clients. Although there were certain predictable and identifiable barriers to success—comparisons to similarly situated individuals were generally informative—one's life circumstances added a layer of nuance that group-level data too often obscured. That is because group-level data is about generalizations rather than specifics.[369] Turning from the general to the specific when deciding a person's fate requires a heavier emphasis on the stories that comprise a person's life.[370]

As with the other potential solutions, this one has its own challenges. Of course, decisionmakers may already take stock of context in a range of ways. That consideration can be based on their personal preferences, biases, or past experiences. But the potential differences in the weights assigned to context and stories by decisionmakers is not unlike what already takes place in the criminal legal system. As long as actors have discretion, they will always be tasked with striking a balance among—and attributing weight to—the information presented to them. That is an unavoidable consequence of being empowered to make decisions about someone else's life.

The answer is to ensure that stories are a part of the decision-making calculus—while qualitative data is used to expose and correct the biased exercise of discretion—rather than in service of making a prediction about someone.[371] Compelling stories about a client's life can shape the outcome of a sentencing proceeding in dramatic ways. Stories can drive judges away from rote sentencing practices and force them to engage facts that "center and humanize" the person to be sentenced and "disrupt judicial inclinations, be they implicit or overt, to base sentences on conclusions derived from bias."[372] Proper consideration of client stories may allow us to move toward a larger ideal: a system of individualized justice, tailored to the circumstances of one's life and weighed against the allegations they face, or the crimes for which they have been convicted.

---

369. Aaron J. Fisher *et al.*, *Lack of Group-to-Individual Generalizability is a Threat to Human Subjects Research*, 115 PROC. OF THE NAT'L ACAD. OF SCI. OF THE U.S E6106, E6113 (2018).

370. One judge's description of sentencing in the context of algorithmic tools is clarifying: "When done correctly, the sentencing process is more art than science. Sentencing requires the application of soft skills and intuitive insights that are not easily defined or even described. Sentencing judges are informed by experience and the adversarial process." Noel L. Hillman, *The Use of Artificial Intelligence in Gauging the Risk of Recidivism*, 58 JUDGES' J. 36, 37 (2019).

371. *See* Roberts, *supra* note 27, at 1727 (suggesting an end to the type of predictive analytics that expand the carceral state).

372. Lindsey Webb, *Slave Narratives and the Sentencing Court*, 42 N.Y.U. REV. L. & SOC. CHANGE 125, 142–43 (2018).

Getting to individualized justice begins with acknowledging that our current system singles out the disfavored among us for control, oppression and punishment, because those on the receiving end are perceived as deserving it.[373] That ideology is cloaked in a historical narrative that refuses to believe that pain even exists at all for people of color and makes it that much easier to punish indefinitely.[374] Combating this mindset requires "tell[ing] a different story," one that allows decisionmakers to "resist the narratives that render" people of color and other marginalized groups "as superhuman to the point of being impervious to pain, and insist[ing] that their pain is our collective responsibility to help heal."[375] Shifting the narrative in that direction means acknowledging the present day impacts of past systems of oppression in to arrive at a set of solutions that stretch beyond those ordinarily deployed by the criminal legal system.[376] Context matters not because it excuses the harms that someone causes, but because it "acknowledges and transforms the realities that made that harm likely."[377]

The current Canadian model of sentencing provides a useful, albeit cautionary, guide.[378] A series of reforms to the Canadian sentencing regime were enacted in 1996.[379] The first Canadian Supreme Court decision that spoke to those reforms, *R. v. Gladue*,[380] deemed the reforms "remedial in nature" and aimed at "ameliorat[ing] the serious problem of overrepresentation of [A]boriginal people in prisons, and . . . encourag[ing] sentencing judges to have recourse to a restorative approach to sentencing."[381] The commitment to act and repair the harms of the past was

373. DANIELLE SERED, UNTIL WE RECKON: VIOLENCE, MASS INCARCERATION, AND THE ROAD TO REPAIR 192–95 (2019).

374. *Id.* at 194–95.

375. *Id.* at 222.

376. *Id.*

377. *Id.* at 224. The type of attention given to context in restorative justice may provide some guidance. As a practice, it requires "acknowledging responsibility for one's actions, acknowledging the impact of one's actions on others, expressing genuine remorse, taking actions to repair the harm to the degree possible, and no longer committing similar harm. . . ." *Id.* at 236–37.

378. There are, of course, notable differences between the structure and character of the Canadian criminal legal system as it pertains to sentencing, chief among them is the incorporation of restorative justice principles in sentencing people of native descent. Toni Williams, *Punishing Women: The Promise and Perils of Contextualized Sentencing for Aboriginal Women in Canada*, 55 CLEV. ST. L. REV. 269, 276–78 (2007); *see also* Anthony N. Doob & Cheryl Marie Webster, *Weathering the Storm? Testing Long-Standing Canadian Sentencing Policy in the Twenty-First Century*, 45 CRIME & JUST. 359, 364–66 (2016) (describing Canadian sentencing practice).

379. *See* Williams, *supra* note 378 at 273–78 (describing sentencing reforms enacted in Canada).

380. [1999] 1 S.C.R. 688 (Can.)

381. *Id.* para. 93. Notably, the parallels between racial disparity in the sentencing of Black people and white people in the United States and Aboriginal people and whites in Canada are striking. The *Gladue* Court took stock of the sentencing disparities in Canada at the time reforms were enacted, pointing out that "[n]ative people come into contact with Canada's correctional

the impetus for the change in the law. According to the Court, "[t]he drastic overrepresentation of [A]boriginal peoples within both the Canadian prison population and the criminal justice system" was expressly recognized as "a sad and pressing social problem" requiring redress in the eyes of the Canadian Parliament.[382]

Gladue traced the sources of sentencing disparity to causes such as "poverty, substance abuse, lack of education, and the lack of employment opportunities for [A]boriginal people . . . . bias against [A]boriginal people and from an unfortunate institutional approach that is more inclined to refuse bail and to impose more and longer prison terms for [A]boriginal offenders."[383] The law placed the onus on sentencing judges to remedy "injustice against [A]boriginal peoples" by requiring judges to "pay particular attention to the circumstances of [A]boriginal offenders, with the implication that those circumstances are significantly different from those of non-[A]boriginal offenders."[384] Those circumstances include:

> The unique systemic or background factors which may have played a part in bringing the particular [A]boriginal offender before the courts; and . . . .[t]he types of sentencing procedures and sanctions which may be appropriate in the circumstances for the offender because of his or her particular [A]boriginal heritage or connection.[385]

Critically, judges must evaluate not only the direct discrimination encountered by native people, but the systemic and institutional structures that drive inequality and injustice.[386]

---

system in numbers grossly disproportionate to their representation in the community . . . . almost 10% of the federal penitentiary population is native (including 13% of the federal women's prisoner population) compared to about 2% of the population nationally . . . ." *Id.* para. 60. At bottom, the Court explained, "[t]he Canadian criminal justice system has failed the Aboriginal peoples of Canada—First Nations, Inuit and Métis people, on-reserve and off-reserve, urban and rural—in all territorial and governmental jurisdictions." *Id.* paras. 60–70. A 2009 statistical analysis reveals that

> Aboriginal persons have consistently comprised 17 to 19% of all adult admissions to Canadian federal penitentiaries for the past decade, even though Indigenous peoples represent only 3% of the Canadian population. The statistics are even more shocking when it comes to admission to provincial jails. In 2007/2008, Indigenous persons comprised 21% of all admissions to provincial jail in Newfoundland and British Columbia, 35% in Alberta, 69% in Manitoba, 76% in the Yukon, 81% in Saskatchewan, and 86% in the Northwest Territories.

David Milward & Debra Parkes, *Gladue: Beyond Myth and Towards Implementation in Manitoba*, 35 MANITOBA L.J. 84, 84–85 (2011).

382. *Gladue*, 1 S.C.R. para. 64.

383. *Id.* para. 65.

384. *Id.* paras. 65–66.

385. *Id.* para. 66.

386. *Id.* paras. 67–69.

The law also directs judges to consider the prevalence of restorative justice in the indigenous community, leading to a regime that requires judges to weigh restorative justice principles in the analysis of an appropriate sentence.[387] In practice, judges may take "judicial notice of the broad systemic and background factors affecting [A]boriginal people, and of the priority given in [A]boriginal cultures to a restorative approach to sentencing."[388] Case-specific details are to come from counsel and a presentence report.[389]

The fact that this regime may yield disparate sentences for Aboriginal people compared to their non-Aboriginal counterparts for the same offense is an accepted function of an individualized system of justice. Such a result is to be expected when a judge undertakes a holistic consideration of the person being sentenced, the person harmed, the affected community, and the available sanctions.[390]

Of course, the gap between theory and practice is often the space where disappointment resides, and the implementation of Canada's sentencing reforms and adherence to *Gladue* is one such space. The system has not dramatically reduced the disparities in sentencing suffered by Aboriginal people compared to their white counterparts in Canada.[391] Yet the failings of *Gladue* and the statutory regime it interpreted can be readily explained. Those explanations can guide the implementation of similar reforms elsewhere.

First, there are natural limits to a sentencing judge's ability to account for and remedy the discrimination and structural barriers to equality that drive people into the criminal legal system.[392] It is also the case that *Gladue* and the law have been applied in an irregular and uncertain fashion across all

387. *Id.* paras. 70–71. *Gladue* described restorative justice:

> as an approach to remedying crime in which it is understood that all things are interrelated and that crime disrupts the harmony which existed prior to its occurrence, or at least which it is felt should exist. The appropriateness of a particular sanction is largely determined by the needs of the victims, and the community, as well as the offender. The focus is on the human beings closely affected by the crime.

Id. para. 71.

388. *Id.* para. 7.

389. *Id.*

390. *Id.* paras. 86–88.

391. *See* Graeme Hamilton, *Twenty Years after Federal Government Changed Sentencing, Aboriginals Still Disproportionately Fill Our Prisons*, NAT'L POST (last updated Aug. 4, 2016), https://nationalpost.com/news/canada/sentence-enough-twenty-years-after-gladue-aboriginals-still-disproportionately-fill-canadian-prisons (describing persistent overrepresentation of indigenous people in Canada's prisons, despite *Gladue*, and noting failure to provide alternatives to incarceration or resources to fully implement sentencing reforms).

392. *R. v. Ipeelee*, [2012] 1 S.C.R. 433, 474–75, para. 69 (Can.).

offenses—most notably those deemed "serious."[393] Everything from the preparation of presentence reports that incorporate the inquiry made by *Gladue* to misgivings about the value of *Gladue* in particular cases have limited its effectiveness.[394] And requiring system actors to be more attentive to context does not, by itself, eliminate the biases they may already harbor in carrying out that mandate.

Notwithstanding the real challenges of implementation, the *Gladue* framework represents a purposeful shift in the information that a sentencing judge is required to focus on and the weight those factors are to be given. Such a shift, if adopted in the United States against the backdrop of algorithmic tools[395] and history of racial inequality, may change our understanding of what justice requires, while avoiding the pitfalls of profiling that can flow from algorithmic tools. One can conceive of a criminal legal system that invokes the use of narratives "to shed light on the conditions of an unjust, racialized institution and to humanize the people placed within it."[396] Numbers can only tell a part of that story.

4. Challenges: Legal and Otherwise.

Many of the recommendations advanced in Part II require a more robust endorsement of race-conscious remedies than the Equal Protection Clause currently allows. The Equal Protection Clause prohibits remedies that "contain[ ] an explicit racial classification [and] laws that assign rights or burdens based on racial classifications"[397] unless they are narrowly tailored to meet a compelling government interest. The types of remedial efforts advanced in Part II, especially those that rest on explicit consideration of race, have run into difficult legal challenges in the past in other contexts[398] and

393. *Id.* para. 86.

394. Milward & Parkes, *supra* note 344, at 86, 94, 96.

395. It is worth noting that algorithmic tools such as risk assessment instruments are widely used in Canada to inform sentencing decisions and issues related to treatment and supervision. *Offender Risk Assessment Practices Vary Across Canada*, PUB. SAFETY CAN., https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/2017-s015/index-en.aspx (last visited Jan. 18, 2020); Simon Fraser University, *Risk Assessment Tools Lead to Fewer Incarcerations Without Jeopardizing Public Safety*, SCIENCEDAILY (Oct. 28, 2019), www.sciencedaily.com/releases/2019/10/191028164344.htm ("Tools are used in nearly every Canadian province and U.S. state, and at least 40 other countries.").

396. Webb, *supra* note 372 at 151. Indeed, such narratives were invoked on the road to the abolition of slavery, as a means to expose the brutality of the system and educate people about the humanity and dignity of people who were enslaved. *Id.*

397. Huq, *supra* note 11, at 1083.

398. *See, e.g.*, *Regents of the Univ. of Cal. v. Bakke*, 438 U.S. 265, 269-271 (1978); *Grutter v. Bollinger*, 539 U.S. 306, 311, 316-17 (2003); *Fisher v. Univ. of Tex.*, 136 S.Ct. 2198, 2205-07 (2016); *see also* Haney-López, *supra* note 296, at 1781–84 (describing how remedial efforts are viewed by the Supreme Court through the lens of "colorblindness" which "consistently imposes the

might here as well.[399] Such challenges are to be expected.[400] The types of obstacles that always accompany the implementation of reforms—from institutional resistance to logistics—warrant attention.[401]

The public appetite for the types of remedial measures and policy recommendations described in Part II also needs to grow. They seek to shift power, redress longstanding harms, and cure significant structural inequities. America has long approached efforts to remedy structural racism as a zero-sum game; where one race or cohort of people may be made whole, the work required to bridge the gap between equity and the status quo necessarily means that another group or race (or several races) must lose out.[402] That zero-sum mentality is complemented by simple racial fear—fear that a status quo which has served the interests of those in power will be upended to their disadvantage.[403] The zero-sum, fear-based worldview has bedeviled remedial measures of all sorts for decades, including efforts to integrate schools, housing and the workforce. The same is true in the criminal legal system. [404]

most stringent form of scrutiny," ensuring that plaintiffs challenging such remedies "virtually always win").

399. There is at least debate about whether differential treatment along racial lines would be barred by Equal Protection jurisprudence. Huq, *supra* note 11, at 1133; *cf.* Deborah Hellman, *Measuring Algorithmic Fairness,* 106 VA. L. REV. 811, 819, 864 (2020) (positing that anti-discrimination law does not pose insurmountable barriers to race conscious remedial efforts aimed at improving the fairness and accuracy of algorithmic tools).

400. The law naturally preserves the status quo. Thus, it "typically works to disadvantage outsiders such as people of color, women, sexual minorities, and the poor." Capers, *supra* note 263, at 24–25.

401. *See, e.g.,* N.Y. STATE SENATE BILL S7506B, 303–12 (2020) https://www.nysenate.gov/legislation/bills/2019/S7506 (expanding the number of crimes eligible for cash bail, three months after the enactment of bail reform); Nick Pinto, *America's Crisis Daddy Andrew Cuomo Exploits Coronavirus Panic to Push Bail Reform Rollback in New York*, INTERCEPT (Mar. 25, 2020) https://theintercept.com/2020/03/25/coronavirus-andrew-cuomo-new-york-bail-reform/ (describing the successful effort to roll back some bail reforms by police, prosecutors, and politicians). *But see* Christopher Robbins, *New York State Legislature Votes to Repeal Law 50-A that Shields Police from Scrutiny*, GOTHAMIST (June 9, 2020) https://gothamist.com/news/new-york-state-legislature-votes-repeal-law-50-shields-police-scrunity (reporting the repeal of 50-A, a law that shielded police misconduct records from the public, which passed despite opposition from police unions state-wide).

402. *See* Michael I. Norton & Samuel R. Sommers, *Whites See Racism as a Zero-Sum Game That They Are Now Losing*, 6 PERSP. ON PSYCH. SCI. 215, 216–17 (2011) (finding that not "only do [w]hites think more progress has been made toward equality than do Blacks, but [w]hites also now believe that this progress is linked to a new inequality—at their expense").

403. Charles M. Blow, *White Extinction Anxiety*, N.Y. TIMES (June 24, 2018), https://www.nytimes.com/2018/06/24/opinion/america-white-extinction.html.

404. Simply put, our criminal legal system is replete with "ill considered policies because we have a pathological political process that caters to the public's fears and emotions without any institutional safeguards or checks for rationality." BARKOW, *supra* note 4, at 12.

Those challenges cannot be dispositive. Resistance to change is natural. Indeed, our racial history has moved through a "reform/retrenchment dialectic."[405] Reforms yield racial progress, only to eventually engender resistance that turns into retrenchment and regress.[406] That is why bold interventions are necessary to upend the status quo.[407] The embedded nature of racism and the status quo preservationist limits of classical liberal reforms means that wholesale change—rather than piecemeal fixes—must be employed.[408] Changes to the law can be driven by cultural shifts, as popular will can shape and reshape legal doctrine.[409] Swaying public sentiment to align the law with the types of remedial measures advanced here would be necessary. Unprecedented national and international mass movements in response to racial injustice and racial hostility[410] have provided some hope about what is possible.[411]

One of Professor Derrick Bell's widely known theories may inform the strategy undertaken to harness the will to foster change: interest convergence. This principle provides that "[t]he interests of blacks in achieving racial equality will be accommodated only when it converges with the interests of whites."[412] In other words, progressive reform has a chance to take hold if the reforms sought are in the interests of the dominant class—those with the power needed to implement them. [413]

Proponents of algorithmic tools—including criminal legal system actors who are aligned with the dominant class—should have an interest in tools that work as advertised to produce a functioning and just criminal legal

405. Carbado, *supra* note 263, at 1607–08.

406. *Id.*

407. "[T]rue change is possible only through radical interventions." Capers, *supra* note 241, at 27.

408. Wing, *supra* note 263, at 52. We must also be prepared to face the fact that even comprehensive changes may not ultimately do the trick. *Id.*

409. *See* BARRY FRIEDMAN, THE WILL OF THE PEOPLE: HOW PUBLIC OPINION HAS INFLUENCED THE SUPREME COURT AND SHAPED THE MEANING OF THE CONSTITUTION 381–84 (2009) (describing how politics and judicial interpretation have shaped Constitutional meaning.).

410. Eduardo Porter, *After the Election, a Nation Tinged with Racial Hostility*, N.Y. TIMES (Nov. 8, 2016), https://www.nytimes.com/2016/11/09/business/after-the-election-a-nation-tinged-with-racial-hostility.html*'*.

411. Larry Buchanan et al., *Black Lives Matter May Be the Largest Movement in U.S. History*, N.Y. TIMES (July 3, 2020), https://www.nytimes.com/interactive/2020/07/03/us/george-floyd-protests-crowd-size.html.

412. Derrick A. Bell, Jr., Brown v. Board of Education *and the Interest-Convergence Dilemma*, 93 HARV. L. REV. 518, 523 (1980).

413. Wing, *supra* note 263, at 48; Carbado, *supra* note 263, at 1608; Capers, *supra* note 263 at 25.

system free from racial inequality.[414] Those who experience the racialized harms and burdens imposed by the criminal legal system likewise have an interest in alleviating those harms. Each of the policy prescriptions in Part II lend themselves to interest convergence because, if implemented, they can best satisfy the reform-oriented concerns of criminal legal system actors and those who have experienced the system as the accused.

Of course, getting people to see how their interests may intersect is not easy. Even a cursory examination of America's political history reveals how hard it is for people to understand how systems of racial oppression can produce harms that are at odds with the fervent sense of superiority that racism foments.[415] Despite these challenges, at the very least, interest convergence sheds some light on how we might overcome the practical hurdles raised by imposing a racial justice lens on the operation of criminal legal system algorithms.

III. ALGORITHMIC TOOLS AND PRAGMATIC ABOLITION

The responses to actuarial risk assessment I have catalogued are largely aimed at shifting the balance of power from actors in the criminal legal system to those who are usually subjected to the harmful treatment by that system. Algorithmic tools may not be able to deliver on the hope that they can reduce racial bias in decision-making. They may not imbue the criminal system with fairness or justice. But their design and use can require a critical inquiry of the way our system operates. And that inquiry produces an opportunity to fundamentally transform our current approach to criminal justice.[416]

I have suggested four interventions: (1) forcing actors to account for the ways in which racism has infected every institution that governs us;[417] (2)

---

414. In New York, for example, District Attorneys in Manhattan, Queens, Bronx, and Brooklyn have shown support for certain parole reform legislation, citing the current law's disproportionate impact on people of color and imploring that "the exorbitant money we are wasting on their reincarceration should be reinvested into programs that make us safer." Darcel Clark *et al.*, *On Parole Violations, Less is More: Three DAs Urge Reform to Stop Sending People Back to Prison*, N.Y. DAILY NEWS, (Mar. 20, 2020) https://www.nydailynews.com/opinion/ny-oped-parole-less-is-more-20200312-bsujoxccpjdh5pocvdgh2d3wny-story.html.

415. Sean Illing, *How the Politics of Racial Resentment Is Killing White People*, VOX (Mar. 19, 2019), https://www.vox.com/2019/3/19/18236247/dying-of-whiteness-trump-politics-jonathan-metzl; Damon Young, *This Is What White Supremacy Looks Like*, NATION (Nov. 9, 2016), https://www.thenation.com/article/archive/this-is-what-white-supremacy-looks-like/.

416. A fundamental transformation refers most readily to the remaking of the American legal system within the frame of racial justice. *See* Paul Butler, *The System is Working the Way It Is Supposed To: The Limits of Criminal Justice Reform*, 104 GEO. L. J. 1419, 1478 (2016) (describing the need to end the current criminal legal system and fundamentally remake it and America).

417. *See supra* Section II.A.

demanding that those who make and deploy algorithmic tools demonstrate that they will not produce racial harms for people of color;[418] (3) holding actors accountable for their decisions by shifting the aim of the tools at those actors;[419] and (4) infusing our system with attention to context, circumstances, and basic dignity.[420] In doing so, the hope is that we can use these algorithmic tools in service of building a different criminal legal system. In this way, the solutions I have outlined operate with a pragmatic abolitionist ethos.

Talk of abolition is sometimes met with derision, because it implies the end of a system without thought given to what comes after the fall.[421] When I advance a pragmatic abolitionist ethos, it does not mean that the criminal legal system as we know it ends immediately, or that abolitionists are not already pragmatic. Practically speaking, a sudden disintegration of America's criminal legal system is not possible.[422] The system is too massive to fall all at once. Instead, a pragmatic abolitionist ethos here means that we press for a wholesale transformation of the criminal legal system while maintaining an "openness to unfinished alternatives,"[423] all while using the tools available to us to do so.

That involves turning to approaches that contradict the premises of the old system, while ensuring those approaches are plausible enough to compete with the system currently in place.[424] Reforms that produce a system that is radically out of step with the status quo will be met with resistance. That is because they seem too far outside the realm of possibility—too unrealistic—given our collective experience with the criminal legal system. But reforms that help to build and shift power in basic ways might be both plausible and effective enough to help carve a path toward transforming the system.[425]

418. *See supra* Section II.B.

419. *See supra* Section II.C.

420. *See supra* Section II.D.

421. *See* ANGELA Y. DAVIS, ARE PRISONS OBSOLETE? 105 (2003) (noting that the question of what replaces jails and prisons following abolition "often interrupts further consideration of the prospects for abolition"). The question misunderstands abolition, because "[a]bolitionists always have their eyes set on a future they are in the process of creating." Dorothy E. Roberts, *Foreword: Abolition Constitutionalism*, 133 HARV. L. REV. 1, 120 (2019) (explaining why abolitionist thinking extends beyond the end of a regime or practice to focus on what comes next).

422. Indeed, "[p]rison abolition is a long-term project that requires strategically working toward the complete elimination of carceral punishment. No abolitionist expects all prison walls to come tumbling down at once." Roberts, *supra* note 421, at 114.

423. Allegra McLeod, *Confronting Criminal Law's Violence: The Possibilities of Unfinished Alternatives*, 8 HARV. UNBOUND 109, 109 (2013).

424. *Id.* at 120.

425. Akbar, *supra* note 343, at 460–69 (detailing an abolitionist ethos built on the need to end "punitive systems of social control" and drive the "reorganization of the state through the redistribution of power and resources").

In the most hopeful view, reformers might use the tools alongside, and in service of, efforts to steadily dismantle the carceral state,[426] with an eye toward ultimately replacing that state with something better suited to delivering justice.[427] That something should be thought of as a "constellation of alternative strategies and institutions"[428] rather than "one single alternative to the existing system of incarceration."[429] Doing so necessarily demands amassing power[430] that drives changes to "unravel rather than widen the net of social control through criminalization."[431] It also means that transformation of the system takes place over time, intermittently, notwithstanding the frustration that will be felt by the deliberate pace of change.

Such a vision is dramatically different than the one offered now in conversations at the intersection of technological tools and the criminal legal system. At present, the turn to tools has—in the best case—been in service of a system that operates more efficiently but retains all of its fundamental characteristics. Indeed, "the state uses artificial intelligence and predictive technologies to reproduce existing inequalities while creating new modes of carceral control and foreclosing imagination of a more democratic future."[432] Yet if we know anything about the criminal legal system—following repeated study, anecdotal evidence, and a wealth of experience—it is not well

426. *See supra* Sections II.C.1 & II.C.2 (discussing potential ways to account for race in algorithmic tools to drive changes in criminal legal system policy).

427. "As movement voices suggest, the abolitionist project is not only negative, it is imaginative; solutions involve social organizations and the reallocation of resources, with investments in jobs, health care, and schools as alternative frameworks for existing investments in policing and incarceration." Akbar, *supra* note 343, at 471. Professor Dorothy Roberts has charted such a course, by suggesting that abolitionists can invoke the U.S. Constitution to advance abolitionist goals. Roberts, *supra* note 421, at 105–20. Specifically, Professor Roberts recommends "using the Constitution to build a society based on principles of freedom, equal humanity, and democracy–a society that has no need for prisons." *Id.* at 110. Such a project is multifaceted. It includes holding courts and legislatures to a faithful reading of the Constitution. *Id.* at 110–13. It also means adopting "nonreformist reforms" that "make transformative changes in carceral systems with the objective of demolishing those systems rather than fixing them." *Id*. at 114. It means mitigating the harms that carceral punishment imposes. *Id*. at 118. And it means using the Constitution to demand the investments required for a society to function without prisons. *Id*. at 119–20.

428. DAVIS, *supra* note 421 at 107.

429. *Id.* at 108.

430. Power here is best defined as "not a thing but rather a capacity composed of active and changing relationships enabling a person, group, or institution to compel others to do things they would not do on their own." Rather than attempting to wrest power away from institutions, we should aim to "make power" through movement building—the result of developing capacities and putting them to work toward transformative ends. RUTH WILSON GILMORE, GOLDEN GULAG: PRISONS, SURPLUS, CRISIS, AND OPPOSITION IN GLOBALIZING CALIFORNIA 247–48 (2007) (emphasis omitted).

431. *Id.* at 242.

432. Roberts, *supra* note 421 at 29.

suited to deliver the justice for which people clamor. It falls short in providing accountability to those harmed by crime, rehabilitation to those who have run afoul of the law, and fairness to all. It traffics in race and inequality. It is the culminating site of all of our social ills. It must be dismantled.[433]

Implementing the types of solutions I have outlined would mean altering the way the system operates while acknowledging and confronting the world as it is. It would mean forcing stakeholders to pay attention to context rather than a forecast, at a societal and individual level, when making decisions about the course of a person's life. It would also mean grappling with the ways that decisions made by criminal legal system actors may be riddled with or reflective of the bias that infects the world around us. This framework represents a reasonable shift away from much of what we currently accept about the criminal legal system—that what someone has done or been accused of renders context irrelevant, that actors cannot be held to account for biased decision-making, and that bias is inevitable and therefore cannot be addressed.

IV. CONCLUSION

Whether or not algorithmic tools will catalyze the types of changes needed to fundamentally alter the way the criminal legal system operates remains to be seen. As communities, scholars, activists, and stakeholders have pointed out, these tools reflect back to us the world that we live in. If we are honest about it, what we see in that reflection is a criminal legal system riddled with racism and injustice. A racial justice lens helps us to understand that and demands that we adjust our responses to what we see to create the type of world that we want to inhabit. We can undertake that work, or continue along our present course and reify the biases and unfairness that already characterize our criminal legal system, but the ultimate choice is ours to make. If we choose wisely, we will use the tools—problems and all—to help us engage in wholesale transformation of the system.

433. "In other words, the way to stop big data's threat to society is not to improve big data. It is to work toward changing the unjust structures that big data supports." Roberts, *supra* note 27, at 1725.