

William & Mary Bill of Rights Journal

Volume 31 (2022-2023)
Issue 2 *The Problem of Mass Incarceration:
Diagnosis and Reform*

Article 7

12-2022

Debiasing Criminal Justice

Sandra Guerra Thompson

Nicole Bremner Cásarez

Follow this and additional works at: <https://scholarship.law.wm.edu/wmborj>



Part of the [Criminal Procedure Commons](#), [Law and Race Commons](#), and the [Law Enforcement and Corrections Commons](#)

Repository Citation

Sandra Guerra Thompson and Nicole Bremner Cásarez, *Debiasing Criminal Justice*, 31 Wm. & Mary Bill Rts. J. 519 (2022), <https://scholarship.law.wm.edu/wmborj/vol31/iss2/7>

Copyright c 2022 by the authors. This article is brought to you by the William & Mary Law School Scholarship Repository.

<https://scholarship.law.wm.edu/wmborj>

DEBIASING CRIMINAL JUSTICE

Sandra Guerra Thompson* and Nicole Bremner Cásarez**

The shame is not in finding that we have unconscious biases or that our current policies have a disproportionate racial impact—the shame lies in refusing to ask the questions and correct the problems.¹

INTRODUCTION

The killing of George Floyd by police officers in Minnesota inspired a summer of protests in 2020, followed by a call for racial reckoning and a professed commitment to reform criminal justice.² Many have condemned the “systemic racism” reflected in countless demographic measures.³ From killings of unarmed men by the police at the front end of the criminal justice system to incarceration rates at the back end, the statistics show stark disparities along racial lines.⁴ These disparities are held up as evidence of racial bias in the system.⁵

* Newell H. Blakely Professor of Law, University of Houston Law Center. Thompson also serves as the Deputy Monitor for the *ODonnell* Consent Decree on misdemeanor bail in Harris County (Houston), Texas.

** Attorney and Emeritus Professor of Communication at the University of St. Thomas in Houston, Texas. Cásarez was one of the lawyers who represented former Texas Death Row inmate Anthony Graves at the time of his exoneration in 2010.

¹ BESIKI KUTATELADZE ET AL., *VERA INST. OF JUST., RACE AND PROSECUTION IN MANHATTAN 1* (2014), <https://www.vera.org/downloads/publications/race-and-prosecution-manhattan-summary.pdf> [<https://perma.cc/C6MT-QBF9>] (quoting Cyrus R. Vance, Jr., former Manhattan District Attorney).

² Following the killing of George Floyd, headlines referred to a summer of racial reckoning which took the form of widespread protests. *See, e.g.*, Ailsa Chang et al., *Summer of Racial Reckoning*, NAT'L PUB. RADIO (Aug. 16, 2020, 9:00 AM), <https://www.npr.org/2020/08/16/902179773/summer-of-racial-reckoning-the-match-lit> [<https://perma.cc/ELP3-RNA5>]. A year later, numerous news articles reflected on the lack of progress after a year of racial reckoning. *See, e.g.*, Judy Woodruff, *Looking Back at a Year-Long Racial Reckoning Since George Floyd's Killing*, PBS NEWS HOUR (May 25, 2021, 6:52 PM), <https://www.pbs.org/newshour/show/looking-back-at-a-year-long-racial-reckoning-since-george-floyds-killing> [<https://perma.cc/NBS2-9Q8R>]; Shereen Marisol Meraji & Gene Demby, *The Racial Reckoning That Wasn't*, NAT'L PUB. RADIO (June 9, 2021, 1:16 AM), <https://www.npr.org/transcripts/1004467239> [<https://perma.cc/BQ82-T9NC>].

³ *See, e.g.*, Radley Balko, Opinion, *There's Overwhelming Evidence That 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-criminal-justice-system/> [<https://perma.cc/5EE7-3M8S>].

⁴ *See infra* note 98 and accompanying text.

⁵ *See infra* note 98 and accompanying text.

Statements about racial bias may be intended as an indictment of a “racist” criminal justice system.⁶ Others have disputed these accusations and argued that the evidence does not support a finding of systemic racism in policing.⁷ Understandably, the law enforcement establishment may hear complaints about racial bias or other forms of bias as a condemnation. If racism is “systemic” then it necessarily implies that the people who operate the system are racists, or at least they act in a manner that has racist results. Others, including law enforcement leaders, counter that any systemic racism does not likely flow from conscious behavior intended to mistreat minorities but from unconscious biases.⁸ A burgeoning literature has addressed implicit racial bias, which refers to biased views that a person may harbor unconsciously.⁹ This research demonstrates that people can act in a racially biased way without realizing it.¹⁰ However, to ascribe disparities in the system to unconscious bias does not much soften the perceived condemnation. To suggest that people act on unconscious racism is still to suggest that they are racists, even if they are unaware of it. To be deemed a racist is to define a person’s character as immoral. For actors in the criminal justice system, it also means that they have acted unlawfully in failing to accord equal protection to all, as well as behaved unethically under professional norms.¹¹ Thus, accusations of racism can be expected to evoke

⁶ See, e.g., Balko, *supra* note 3.

⁷ See, e.g., Heather Mac Donald, Opinion, *The Myth of Systemic Police Racism*, WALL ST. J. (June 2, 2020, 1:44 PM), <https://www.wsj.com/articles/the-myth-of-systemic-police-racism-11591119883> [<https://perma.cc/62MM-LR3Z>] (arguing that police shootings are a function of violent police encounters and that the commission of violent crime by Black people explains police shootings and also arguing that more Black and Hispanic police officers are responsible for shootings than white officers).

⁸ The International Association of Chiefs of Police “Bias-Free Policing Policy” explains that “racism” refers to an explicit bias or prejudice and that differential treatment likely occurs as a result of implicit or unconscious biases, as opposed to racism. L. ENF’T POL’Y CTR., INT’L ASS’N OF CHIEFS OF POLICE, BIAS-FREE POLICING 6 (2021), <https://www.theiacp.org/sites/default/files/2021-01/Bias-Free%20Policing%202021-01.pdf> [<https://perma.cc/3VN4-PBDS>].

⁹ See *infra* Part II and Section II.A.

¹⁰ See JENNIFER K. ELEK & ANDREA L. MILLER, NAT’L CTR. FOR STATE CTS., THE EVOLVING SCIENCE ON IMPLICIT BIAS 6–7 (2021) [hereinafter NCSC REPORT], <https://ncsc.contentdm.oclc.org/digital/collection/accessfair/id/911> [<https://perma.cc/Y22R-VPQW>].

¹¹ See U.S. CONST. amend. XIV, § 1. Equal protection and other constitutional violations can make a state actor subject to civil and criminal penalties under federal law. Section 1983 of the Civil Rights Act of 1871 allows injured individuals to sue public officials for violations of their constitutional rights. See 42 U.S.C. § 1983. Prosecutors can also file charges against public officials for civil rights violations under 18 U.S.C. § 242.

Various professional standards govern policing and prosecution. The Code of Ethics of the International Association of Chiefs of Police, for example, states, “I will never act officiously or permit personal feelings, prejudices, political beliefs, aspirations, animosities or friendships to influence my decisions.” See *Law Enforcement Code of Ethics*, INT’L ASS’N OF CHIEFS OF POLICE, <https://www.theiacp.org/resources/law-enforcement-code-of-ethics> [<https://perma.cc/8ZL4-CWH6>] (last visited Dec. 8, 2022). For prosecutors, the National

forceful denials, resentment, and a resultant unwillingness to acknowledge a need for reform.¹²

This Article calls for a change in perspective. Rather than view bias as an accusation or moral flaw, criminal justice officials should instead consider it as scientists do—as an inevitable source of error to be minimized so as to produce outcomes that better reflect objective truth. Many types of bias can threaten the validity of basic scientific research, such as drug testing, and biases can introduce unwarranted disparities in the clinical practice of medicine as well.¹³ Scientists strive to identify possible sources of bias and then to find demonstrably effective methods for reducing those biases.¹⁴ Minimizing bias in scientific endeavors is simply a form of quality control.¹⁵ The practice of medicine also provides a highly relevant parallel for the legal profession in that wide racial and ethnic disparities have plagued the health care system, such as in the area of pain treatment.¹⁶ This Article explores the data

District Attorneys Association's National Prosecution Standards states that, in screening cases to decide whether to file a charge, a prosecutor should not consider the "[c]haracteristics of the accused that have been recognized as the basis for invidious discrimination, insofar as those factors are not pertinent to the elements or motive of the crime." NAT'L DIST. ATT'YS ASS'N, NATIONAL PROSECUTION STANDARDS § 4-1.4 (3d ed. 2009), <https://ndaa.org/wp-content/uploads/NDAA-NPS-3rd-Ed.-w-Revised-Commentary.pdf> [<https://perma.cc/P9AH-C777>]. Likewise, the American Bar Association's Criminal Justice Standards states that a prosecutor "should not manifest or exercise, by words or conduct, bias or prejudice based upon race, sex, religion, national origin, disability, age, sexual orientation, gender identity, or socioeconomic status." AM. BAR ASS'N, CRIMINAL JUSTICE STANDARDS FOR THE PROSECUTION FUNCTION 3-1.6 (4th ed. 2017).

¹² The "Blue Lives Matter" movement seems to have organized as a countermeasure to support law enforcement interests in the face of the "Black Lives Matter" movement. The stated mission of the #BlackLivesMatter group includes "to eradicate white supremacy and build local power to intervene in violence inflicted on Black communities by the state and vigilantes." *See About*, BLACK LIVES MATTER, <https://blacklivesmatter.com/about/> [<https://perma.cc/K75M-LEFF>] (last visited Dec. 8, 2022). The Facebook page for Blue Lives Matter shows a "Thin Blue Line" flag with the caption, "Sometimes there's justice. Sometimes there's just us." *Blue Lives Matter*, FACEBOOK, <https://www.facebook.com/bluematters/> [<https://perma.cc/NZA9-F4ZM>] (last visited Dec. 8, 2022). In particular, the "Thin Blue Line" flag and insignia have been seen as symbols of support for law enforcement in the face of accusations of systemic racism, but some say extremists hold it up as a symbol of white supremacy. *See, e.g.*, Janelle Griffith, *Police Chief Bans 'Thin Blue Line' Imagery, Says It's Been 'Co-Opted' by Extremists*, NBC NEWS (Jan. 29, 2021, 5:09 PM), <https://www.nbcnews.com/news/us-news/police-chief-bans-thin-blue-line-imagery-says-it-s-n1256217> [<https://perma.cc/6HYD-8UDN>].

¹³ *See infra* notes 46–47 and accompanying text.

¹⁴ *See* Christopher J. Pannucci & Edwin G. Wilkins, *Identifying and Avoiding Bias in Research*, 126 PLASTIC & RECONSTRUCTIVE SURGERY 619, 619–25 (2010).

¹⁵ *See infra* notes 94, 110 and accompanying text.

¹⁶ *See* Salimah H. Meghani et al., *Advancing a National Agenda to Eliminate Disparities in Pain Care: Directions for Health Policy, Education, Practice, and Research*, 13 PAIN MED. 5, 5–6 (2012).

on racial and ethnic disparities and the various interventions recommended both for medical professionals at the micro level as well as systemic reforms at the macro level.

Next, this Article highlights recent guidance from the National Council of State Courts regarding those intervention techniques that it found to be most effective in reducing implicit bias in criminal justice.¹⁷ We also briefly summarize proposals made by legal scholars calling for the use of blinding and other techniques at various key points in the criminal justice process.¹⁸ Finally, this Article describes the efforts of two California District Attorneys' offices to employ artificial intelligence blinding technology to reduce implicit bias in prosecutorial charging decisions.¹⁹ Our survey of the studies and pilot projects applying scientifically developed strategies for reducing implicit bias demonstrates that leaders in the criminal justice arena have only begun the search for effective remedies, but much work remains to be done if we ever hope to reduce unwarranted disparities and achieve true justice.

I. VIEWING BIAS THROUGH THE LENS OF SCIENCE: BIAS REDUCTION AS QUALITY CONTROL

The Supreme Court has written about the scientific method in the seminal case of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*²⁰ The Court wrote: “[I]n order to qualify as ‘scientific knowledge,’ an inference or assertion must be derived by the scientific method.”²¹ The scientific method tests the integrity of a hypothesis in several ways through a process known as “validation.”²² The validation process ensures that expert conclusions constitute true scientific knowledge.²³ Thus, scientists must take into account any aspect of the process that might introduce the possibility of error.²⁴ Indeed, an important aspect of good scientific work is to determine the “known or potential rate of error” for any process that produces an objective finding.²⁵

Bias tends to introduce error into a decision-making process.²⁶ To reduce the incidence of error in a scientific process, scientists strive to discern and control

¹⁷ See *infra* notes 100–15 and accompanying text.

¹⁸ See *infra* notes 118–25 and accompanying text.

¹⁹ See *infra* Section II.B.

²⁰ *Daubert v. Merrell Dow Pharms.*, 509 U.S. 579, 590 (1993).

²¹ *Id.*

²² *Id.*

²³ *Id.* The Court incorporated the principal features of the scientific method as part of its guidance for the lower courts to follow in judging the reliability of scientific evidence. *Id.* at 593–94.

²⁴ One feature of the scientific method involves consideration of any known error rate. *Id.* at 594.

²⁵ *Id.*

²⁶ See Pannucci & Wilkins, *supra* note 14, at 619.

sources of possible bias.²⁷ While there are many definitions of bias in the literature, one medical article defines a bias as “any tendency which prevents unprejudiced consideration of a question.”²⁸ Further, it explains that bias causes “systematic error . . . by selecting or encouraging one outcome or answer over others.”²⁹ Rather than concern themselves with whether bias may exist, medical researchers understand that “some degree of bias is nearly always present.”³⁰ Nor do scientists believe that bias in medical research, for example, is ordinarily a product of malicious intent.³¹ Rather, they appreciate that “[e]xplicit bias [which] requires that a person is aware of his/her [prejudiced] evaluation of a group . . . has declined significantly over the past 50 years and is now considered unacceptable in general society.”³² On the other hand, “implicit bias appears to be common and persistent.”³³ In this section, we highlight a few lessons about the way that the medical profession approaches the study of bias, and we address some of the techniques used for minimizing it.

Scientists realize that addressing bias must involve a layered approach. First, many types of bias will affect the operation of a system or scientific process at numerous points from start to finish.³⁴ For example, one scientific organization at Oxford University devotes itself strictly to creating an annotated catalogue of known biases affecting health studies.³⁵ The existence of so many types of bias that can occur at various points in a system means that researchers will need to use several remedial strategies to address these different types of bias.³⁶

Moreover, one remedial strategy will often necessitate another to reinforce the efficacy of its implementation. Thus, effective strategies for bias reduction involve a commitment to addressing the possibility of biases at various points in the system using multiple, reinforcing remedial approaches, rather than seeking a simple, one-step fix.³⁷ Within a process or system, there will be numerous points at which

²⁷ *Id.*

²⁸ *Id.*

²⁹ *Id.* (quoting *Bias*, MERRIAM-WEBSTER.COM, <https://www.merriam-webster.com/dictionary/bias> [<https://perma.cc/5B6W-TAQE>] (last visited Dec. 8, 2022)).

³⁰ *Id.*

³¹ See Irene V. Blair et al., *Unconscious (Implicit) Bias and Health Disparities: Where Do We Go from Here?*, 15 PERMANENTE J. 71, 71 (2011).

³² *Id.* (citation omitted).

³³ *Id.* (citations omitted).

³⁴ See Pannucci & Wilkins, *supra* note 14, at 619 (“Bias can occur at any phase of research, including study design or data collection, and in the process of data analysis and publication.”).

³⁵ See *About*, UNIV. OF OXFORD CTR. FOR EVIDENCE-BASED MEDICINE, CATALOGUE OF BIAS, <https://catalogofbias.org/about/> [<https://perma.cc/9TN7-9WV4>] (last visited Dec. 8, 2022).

³⁶ See Pannucci & Wilkins, *supra* note 14, at 621 tbl.1 (listing numerous types of bias, such as selection bias, channeling bias, interviewer bias, chronology bias, recall bias, transfer bias, performance bias, and citation bias).

³⁷ *Id.* (noting numerous types of interventions to avoid different types of bias at different phases of a clinical trial).

participants will gather and document information and make various decisions.³⁸ At every point at which a human being exercises decision-making discretion, even in simply how to gather and record information, a choice may be influenced by some type of bias.³⁹ In a complex process or system, numerous types of bias will affect the process from the design of the process to its implementation and the final assessment of outcomes.⁴⁰

To review some types of remedial strategies, consider the randomized control trial (RCT). Medical researchers consider RCTs as the gold standard for unbiased research.⁴¹ Using RCTs, researchers can study the effects of a medical intervention by comparing two similar groups of people to see whether those receiving the medical intervention experience some benefit.⁴² Ensuring that the two groups of people are similar in all important respects is crucial to the accuracy of the study's results.⁴³ If bias-reduction techniques are not used in the design of the study, the results can be biased from the outset by "selection bias," which happens "when the criteria used to recruit and enroll patients into separate study cohorts are inherently different."⁴⁴ Selection bias will result in the two comparison groups being different from the start, thus invalidating any comparison of how the medical intervention affects the two groups.⁴⁵ To minimize the effect of selection bias, researchers use randomization, referring to a random selection process for determining comparison cohorts.⁴⁶

Once the two groups are randomly selected, bias can nonetheless affect the accuracy of the results if either the subjects or researchers know which cohort receives the medical intervention and which the placebo. Subjects who know which group they belong to may act differently, and researchers may treat them differently as well, thus biasing the outcome assessments.⁴⁷ To reduce, if not eliminate, this type

³⁸ See *id.* at 619.

³⁹ *Id.*

⁴⁰ *Id.*

⁴¹ Eduardo Hariton & Joseph J. Locascio, *Randomised Controlled Trials—the Gold Standard for Effectiveness Research*, 125 *BJOG* 1716, 1716 (2018) (advocating for RCTs as a study design to reduce bias).

⁴² *Id.*

⁴³ *Id.*

⁴⁴ Pannucci & Wilkins, *supra* note 14, at 620–21.

⁴⁵ *Id.*

⁴⁶ Paul J. Karanicolas et al., *Blinding: Who, What, When, Why, How?*, 53 *CANADIAN J. SURGERY* 345, 345 (2010); Hariton & Locascio, *supra* note 41, at 1716 ("[R]andomisation reduces bias and provides a rigorous tool to examine cause-effect relationships between an intervention and outcome.").

⁴⁷ See Karanicolas et al., *supra* note 46, at 345 (mentioning that, following random selection of cohorts, blinding is used in RCTs to "prevent subsequent differential cointerventions or biased assessment of outcomes"); see also Dorothy Forbes, *Blinding: An Essential Component in Decreasing Risk of Bias in Experimental Designs*, 16 *EVIDENCE BASED NURSING* 70, 70 (2013) ("Lack of blinding in randomised controlled trials (RCTs) has been

of bias, researchers use a “double-blind” process that prevents both subjects and researchers from knowing which group is receiving the tested intervention.⁴⁸ Double-blind practices have been encouraged, even for clinical trials involving surgical interventions.⁴⁹ The practice of blinding (or masking) is considered “a critical . . . feature of RCTs.”⁵⁰ Even for the evaluation of the *findings* from clinical studies, scientists employ similar techniques for reducing biases in the assessments of outcomes. For example, with regard to the safety of medical imaging products (e.g., iodinated compounds used in radiography and CT), the Food and Drug Administration (FDA) recommends fully blinded evaluations of the images produced in clinical drug trials, and the FDA recommends that the sequence of assessments be randomized.⁵¹

As part of an RCT (as well as in the medical clinical setting), medical personnel typically ask questions of the subjects in the study (or patients in a clinical setting).⁵² Medical professionals understand that “interviewer bias” can affect the answers given by the people being questioned, depending on how the questions are phrased, or how the answers are recorded or interpreted.⁵³ To address interviewer bias, researchers have identified two important strategies that should generally be used in combination. First, the amount of discretion exercised by an interviewer can be reduced through the use of standardized protocols and objective guidelines.⁵⁴

shown to be associated with more exaggerated estimated intervention effects, by nine percent on average.”).

⁴⁸ See Karanicolas et al., *supra* note 46, at 346.

⁴⁹ Karolina Wartolowska et al., *Blinding in Trials of Interventional Procedures Is Possible and Worthwhile* 1, 3, 6 (Jan. 30, 2018) (on file with F1000Research), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5717470/pdf/f1000research-6-14943.pdf> [<https://perma.cc/B9FG-KG9D>].

⁵⁰ See Karanicolas et al., *supra* note 46, at 345; see also Forbes, *supra* note 47, at 70 (noting that blinding is also known as “masking”).

⁵¹ U.S. FOOD & DRUG ADMIN., GUIDANCE FOR INDUSTRY: DEVELOPING MEDICAL IMAGING DRUG AND BIOLOGICAL PRODUCTS, PART 3: DESIGN, ANALYSIS, AND INTERPRETATION OF CLINICAL STUDIES 12–14 (2004). The term “medical imaging agents” is defined in great detail in U.S. FOOD AND DRUG ADMINISTRATION, GUIDANCE FOR INDUSTRY: DEVELOPING MEDICAL IMAGING DRUG AND BIOLOGICAL PRODUCTS, PART 1: CONDUCTING SAFETY ASSESSMENTS 2 (2004). The FDA guidance also recommends a process known as “sequential unblinding” or “sequential unmasking” for the evaluation process. This involves conducting evaluations of imaging products fully blinded to the pertinent clinical information so as to remove biases and then providing progressively more clinical information to the test agent for purposes of comparison. U.S. FOOD & DRUG ADMIN., GUIDANCE FOR INDUSTRY, PART 3, *supra* note 51, at 11–13. A full discussion of sequential unmasking is beyond the scope of this Article, however, one of us has previously written about the potential use of this process within crime laboratories. See SANDRA GUERRA THOMPSON, COPS IN LAB COATS: CURBING WRONGFUL CONVICTIONS THROUGH INDEPENDENT FORENSIC LABORATORIES 141–43 (2015).

⁵² Pannucci & Wilkins, *supra* note 14, at 620–21.

⁵³ *Id.* at 621 (defining “interviewer bias” as “a systematic difference between how information is solicited, recorded, or interpreted”).

⁵⁴ *Id.* at 620–21.

Second, this approach can be enhanced by training practitioners on how to best gather, record and interpret patient data to minimize “interobserver variability when multiple individuals are gathering and entering data.”⁵⁵ Targeted training programs for the collection and gathering of data using standardized protocols have proved effective in reducing interviewer bias.⁵⁶ But researchers encourage the use of blinding in combination with these interventions, or, if blinding is not possible, having different examiners measure the outcome.⁵⁷ This type of interviewer training is to be distinguished from general implicit bias training used as a sole remedy, which may have only limited efficacy in the short term and may even be counterproductive.⁵⁸

While scientists have successfully implemented bias reduction techniques for RCTs, physicians have had greater difficulty implementing effective bias reduction techniques in clinical practice. In a remarkable parallel to the challenges faced in the criminal justice system, the medical profession has struggled to find solutions for serious racial and ethnic disparities in the delivery of health care.⁵⁹ Congress in 2000 passed the Minority Health and Health Disparities Research and Education Act to promote research on minority health disparities.⁶⁰ Then in 2003, the Institute of Medicine published an influential report, *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*, detailing the wide disparities in services and health outcomes, even after controlling for income, neighborhood, comorbid illnesses, and health insurance type.⁶¹ The research shows disparities across many areas of health care from treatment in cardiovascular care to cancer care and maternal morbidity.⁶² Researchers have also found bias and disparities in the provision of health care based on gender.⁶³

⁵⁵ *Id.* at 620.

⁵⁶ *Id.*

⁵⁷ *Id.*

⁵⁸ See Blair et al., *supra* note 31, at 75 (“[A] direct approach of confronting an individual with evidence of bias may actually have little effect on that bias [and] . . . research shows that intentionally trying to suppress bias may actually make it ‘rebound’ at a later time.”); see also *infra* notes 112–15 and accompanying text.

⁵⁹ See Salimah H. Meghani et al., *Advancing a National Agenda to Eliminate Disparities in Pain Care: Directions for Health Policy, Education, Practice, and Research*, 13 PAIN MED. 5, 5 (2012).

⁶⁰ Minority Health and Health Disparities Research and Education Act of 2000, Pub. L. No. 106-525, 114 Stat. 2495; see also Meghani et al., *supra* note 59, at 5.

⁶¹ See generally INST. OF MED., *UNEQUAL TREATMENT: CONFRONTING RACIAL AND ETHNIC DISPARITIES IN HEALTH CARE* (Brian D. Smedley, Adrienne Y. Stith & Alan R. Nelson eds., 2003).

⁶² See, e.g., *id.* at 5–6; Martha Hostetter & Sarah Klein, *In Focus: Reducing Racial Disparities in Health Care by Confronting Racism*, COMMONWEALTH FUND (Sept. 27, 2018), <https://www.commonwealthfund.org/publications/2018/sep/focus-reducing-racial-disparities-health-care-confronting-racism> [<https://perma.cc/EA7H-ZK3V>] (addressing disparities in cancer treatment and maternal morbidity).

⁶³ See, e.g., Jess Severson, *Taking the Bias Out of Pain Diagnosis*, SCI. AM.: VOICES

Pain treatment has received special attention as the disparities are pronounced and the negative effects for Black patients and other ethnic minority groups are far-ranging, including “disability, poor quality of life, relational problems, lost income and productivity, and higher health care utilization.”⁶⁴ One meta-analysis of the empirical literature on emergency department use of pain treatments found Black patients were thirty-six to forty percent less likely to receive any analgesia compared to white patients and only slightly lower disparities in the receipt of opioids for acute pain.⁶⁵ Hispanic patients were twenty-five to thirty percent less likely than white patients to receive any analgesia and slightly less likely to receive opioids for acute pain.⁶⁶

Pain measurement is inherently subjective.⁶⁷ Physicians exercise wide discretion in assessing pain, and base their conclusions on many subjective factors expressed in different ways by individual patients.⁶⁸ Typically, doctors ask patients to rate their pain on a scale from one to ten.⁶⁹ However, even similar numerical measurements of pain might result in different treatment depending on the patient.⁷⁰ This wholly subjective assessment based on communication with patients allows for the introduction of implicit bias based on stereotypes and also due to cultural communication barriers.⁷¹ Moreover, patients who perceive that a physician views them in stereotypical ways may experience “stereotype threat,” which “is a stressful psychological state

(July 16, 2019), <https://blogs.scientificamerican.com/voices/taking-the-bias-out-of-pain-diagnosis/> [<https://perma.cc/2GUK-55CX>] (noting a study finding that women wait an average of sixty-five minutes in the ER, compared to men’s forty-nine minutes and another finding different types of pain relief medications prescribed to women as compared to men).

⁶⁴ See Meghani et al., *supra* note 59, at 6.

⁶⁵ *Id.* at 8.

⁶⁶ See Paulyne Lee et al., *Racial and Ethnic Disparities in the Management of Acute Pain in US Emergency Departments: Meta-Analysis and Systematic Review*, 37 AM. J. EMERGENCY MED. 1770, 1772 (2019) (finding racial and ethnic disparities in analgesic use based upon a review of seventeen studies of pain and analgesic administration).

⁶⁷ See Janice A. Sabin, *How We Fail Black Patients in Pain*, ASS’N AM. MED. COLLS. (Jan. 6, 2020), <https://www.aamc.org/news-insights/how-we-fail-black-patients-pain> [<https://perma.cc/JXZ2-CXAW>] (noting that medical providers rely more on their personal judgment in cases of ambiguity, such as when pain is not caused by an obvious physical condition, and that meta-analysis has shown the greatest racial disparity in pain treatment for conditions like backache, migraine, and abdominal pain).

⁶⁸ *Id.*

⁶⁹ See Severson, *supra* note 63, at 2.

⁷⁰ *Id.*

⁷¹ See Sabin, *supra* note 67 (noting a study that forty percent of first- and second-year medical students endorsed the false belief that “[B]lack people’s skin is thicker than white people’s,” making them less sensitive to pain, as well as the stereotype frequently portrayed in the media associating Black people with substance abuse); Louis A. Penner et al., *Reducing Racial Health Care Disparities: A Social Psychological Analysis*, 1 POL’Y INSIGHTS FROM BEHAV. & BRAIN SCIS. 204, 206 (explaining that when physicians exhibit implicit bias, Black patients can be sensitive to these behaviors and will report less patient-centered care and lower levels of trust and satisfaction).

that occurs when a person fears being judged by others on the basis of negative stereotypes.”⁷² Inducing this psychological state in patients “may impair patient-clinician communication, reduce self-efficacy, and increase mistrust” which can lead to less functional behavior by both patients and physicians.⁷³

The impact of implicit biases may have the greatest effect on patients who are both Black and female. For example, in studies of maternal morbidity, Black women die from pregnancy-related complications at three to four times the rate of white women, and education, socioeconomic status, and other factors do not protect Black women from the risk, nor do smoking, drug abuse, and obesity explain the differences.⁷⁴ These findings have led some to conclude that “being a [B]lack woman in America is, itself, a risk factor.”⁷⁵ Dr. Neel Shah, assistant professor of obstetrics, gynecology and reproductive biology at Harvard Medical School, notes that “the studies suggest . . . that we believe [B]lack women less when they express symptoms, and we tend to undervalue their pain.”⁷⁶

Researchers have proposed a range of techniques to address the concerns about implicit bias and health outcome disparities. Following the Institute of Medicine 2003 report, medical school training and postgraduate medical education added courses on racial health disparities into the curriculum in many medical schools.⁷⁷ However, training alone has proved insufficient, as subsequent studies have shown. Coursework has only limited impact in affecting physician behavior, especially because the information about disparities may not help physicians to recognize how their own personal beliefs and actions contribute to these disparities.⁷⁸ Furthermore, attempts to change a person’s implicit biases, which are “habits of mind acquired over time,” can be difficult to achieve and can result in “backlash against external efforts to change how they think.”⁷⁹

Nonetheless, some see the potential to reduce implicit bias through training that encourages physicians to use patient-centered communication in which the physician

⁷² See Blair et al., *supra* note 31, at 75.

⁷³ *Id.*

⁷⁴ See Hostetter & Klein, *supra* note 62.

⁷⁵ *Id.*

⁷⁶ *Id.* In the legal academy, Kimberle Crenshaw has also urged the consideration of the special issues that arise in the “intersectionality” of gender and race. See, e.g., Kimberle Crenshaw, *Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color*, 43 STAN. L. REV. 1241, 1242 (1991).

⁷⁷ See Penner et al., *supra* note 71, at 208; Sabin, *supra* note 67.

⁷⁸ See Penner et al., *supra* note 71, at 208.

⁷⁹ *Id.*; see also Brian B. Drwecki, *Education to Identify and Combat Racial Bias in Pain Treatment*, 17 AMA J. ETHICS 221, 221–22 (2015) (“[F]eedback indicating that one’s responses may be racially biased causes guilt and self-criticism and in some cases creates withdrawal motivations that make people less likely to confront and ameliorate their racial prejudice,” and an individual’s “natural efforts at prejudice reduction (i.e., exerting willpower to suppress biased impulses) are not effective and can lead to greater expression of prejudice in the long run.”).

views each patient as a unique individual rather than mainly as a representative of a racial or social group.⁸⁰ A field study had great success in asking patients about their racial and cultural identities and whether they had experienced racism in health care.⁸¹ The doctors found it extremely helpful to frame the issue as something that has happened to other people, so as to make patients feel more comfortable speaking up.⁸² These efforts promote effective physician-patient communication, which is essential to subjective assessments based on self-reporting.⁸³

Several researchers have called for the use of guidelines (including standardized treatment criteria and procedures) and standardized checklists as a way to remove individual discretion, and therefore to minimize the erroneous decisions produced from biased judgment.⁸⁴ For pain treatment, the Centers for Disease Control and Prevention (CDC) offers a *Guideline for Prescribing Opioids for Chronic Pain* that cover several topics, and the American Academy of Pain Medicine has published guidance for implementing the CDC's guideline through standardized protocols.⁸⁵ Additionally, medical school education can include effective strategies for reducing implicit bias in patient care, including the use of clinical guidelines and standardized checklists.⁸⁶ In the medical field, proposing standardized treatment criteria and procedures runs counter to a patient-centered approach that encourages individuation of patients, and thus encounters some resistance.⁸⁷ One group of researchers conclude that “[n]evertheless, a reasonable adoption of standardized care does have potential for reducing the racial disparities in health care, while improving the quality of care for everyone.”⁸⁸

⁸⁰ Penner et al., *supra* note 71, at 208–09. The authors also found that using messages and symbols that repeatedly stressed the team nature of the physician-patient interaction resulted in increasing the level of trust patients experienced and their willingness to adhere to the physician's recommendations. *Id.* at 209–10. Laboratory testing has also found success in “reduc[ing] stereotyping and engender[ing] empathy, understanding, and perceptions of equality” through “equal-status contact, which involves members from different racial and ethnic groups interacting as equals in situations of shared power.” Drwecki, *supra* note 79, at 223. Moreover, training that provides “exposure to counter-stereotypic African American exemplars, competing on teams with members of other groups, and repeatedly practicing associating positive words with members of other racial groups were all shown to successfully reduce nonconscious biases, across multiple labs.” *Id.*

⁸¹ See Hostetter & Klein, *supra* note 62.

⁸² *Id.*

⁸³ See Sevetson, *supra* note 63.

⁸⁴ See Sabin, *supra* note 67; Penner et al., *supra* note 71, at 210 (stating that “[n]oted surgeon and writer Atul Gawande . . . proposed . . . standardiz[ed] treatment criteria and procedures, reducing unnecessary variability in medical treatments” and that his proposal was “controversial but innovative” and “would represent a radical departure from how medical care is currently provided.” (emphasis omitted)).

⁸⁵ See Sabin, *supra* note 67.

⁸⁶ *Id.*

⁸⁷ See Penner et al., *supra* note 71, at 210.

⁸⁸ *Id.*

Interestingly, two recent breakthroughs in new technologies may provide objective methods for physicians to measure pain which may facilitate the consistent application of pain management guidelines to promote equal treatment.⁸⁹ One method involves the measurement of brain waves to rate pain, and the other relies on blood testing.⁹⁰ As is generally true, objective determinations, not influenced by implicit biases, can ameliorate not only racial and ethnic biases, but also other types such as gender-based bias.⁹¹

Two other proposed measures bear consideration. First, numerous medical researchers advocate for the collection of data and further research on bias as a means of developing evidence-based interventions for medical education and medical practice to combat implicit bias and disparities.⁹² Second, the use of standardized treatment and data collection also permits “oversight by individuals with organizational responsibility for quality of service.”⁹³ These two proposals operating at the same time—standardization combined with data collection and analysis—would enable quality control oversight by organizational managers and the development of other effective interventions.⁹⁴

In summary, this review of the medical literature demonstrates that the medical profession has engaged in a continual effort to study and address numerous types of bias in both medical research and clinical practice. The use of randomization and double-blind techniques in RCTs, supported by training and standardized guidelines, is widely utilized as a means of minimizing numerous types of possible bias in medical research.⁹⁵ In clinical practice, the effort to reduce racial, ethnic and gender disparities continues as researchers search for interventions that can provide more objective or standardized practices to reduce the untoward influences of implicit bias.⁹⁶ These commendable efforts to study, track and combat implicit bias offer many lessons for the criminal justice system, which has only recently begun its own search for effective interventions to reduce bias.

⁸⁹ See Severson, *supra* note 63.

⁹⁰ *Id.*; Sabin, *supra* note 67.

⁹¹ Sabin, *supra* note 67.

⁹² See, e.g., Hostetter & Klein, *supra* note 62 (advocating that states require health care providers to track racial and ethnic disparities in treatment for a wide range of conditions and making this data public); Penner et al., *supra* note 71, at 210 (advocating for the aggregation of information within large health care systems to make any racial disparities apparent and to promote more effective, evidence-based interventions); Sabin, *supra* note 67 (calling for the collection of data on pain management by patients’ race and ethnicity, as well as gender, insurance status and other significant characteristics, so as to enable creating strategies to address disparities); Drwecki, *supra* note 79, at 224 (advocating for medical schools to test for racial bias in its students, and, if biases are found, to use the data to develop effective interventions).

⁹³ See Penner et al., *supra* note 71, at 210.

⁹⁴ *Id.*

⁹⁵ See, e.g., Hariton & Locascio, *supra* note 41, at 1716.

⁹⁶ *Id.*

II. BIAS REDUCTION STRATEGIES FOR THE CRIMINAL JUSTICE SYSTEM

Like the practice of medicine, as well as the process used for clinical trials in medical research, the criminal justice system can be viewed as a system involving the gathering of information and human decision-making by multiple people from start to finish. As such, like any other information gathering and decision-making process, the criminal justice system is affected by many types of implicit bias.⁹⁷ Scholars have explored many types of bias in the criminal justice system that result in unwarranted disparities in outcomes.⁹⁸ Moreover, recent events and calls for racial reckoning urge criminal justice leaders to renew their efforts to curb the effects of implicit bias that result in unwarranted racial and ethnic disparities.⁹⁹

A. The Research on Implicit Bias and Criminal Justice

As part of the search for remedial interventions to reduce implicit bias in the legal profession, a 2021 report by the National Council of State Courts (NCSC) provides useful guidance for criminal justice leaders based on extensive scientific research on implicit biases and the remedial approaches to combat them.¹⁰⁰ The NCSC

⁹⁷ See NCSC REPORT, *supra* note 10, at 6–7 (referencing studies showing the effects of judges’ implicit biases on sentencing decisions, police officers’ implicit biases on decisions to shoot criminal suspects of different races, and trial court judges’ decisions affected by race and gender).

⁹⁸ See, e.g., Margaret Bull Kovera, *Racial Disparities in the Criminal Justice System: Prevalence, Causes, and a Search for Solutions*, 75 J. SOC. ISSUES 1139, 1140 (2019) (addressing racial disparities in policing, prison populations, and jury service); Marvin D. Free, Jr., *Racial Bias and the American Criminal Justice System: Race and Presentencing Revisited*, 10 CRITICAL CRIMINOLOGY 195, 197 (2001) (critiquing fifty-two studies on the effect of race on bail, pretrial release, and the dismiss/reject/prosecute decision); Cynthia E. Jones, “Give Us Free”: *Addressing Racial Disparities in Bail Determinations*, 16 N.Y.U. J. LEGIS. & PUB. POL’Y 919, 921–22 (2013) (exploring racial disparities in the pretrial bail determination and reporting on policy changes implemented in Saint Louis County (Duluth), Minnesota); STACEY J. BOSICK, UNIV. OF DENVER, RACIAL DISPARITIES IN PROSECUTORIAL OUTCOMES 1–4 (2021), https://www.denverda.org/wp-content/uploads/news-release/2021/Racial-Disparities-in-Prosecutorial-Outcomes_March2021_final-002.pdf [<https://perma.cc/3EU4-9XAB>] (studying Denver County District Attorney’s data and finding disparities in dismissals, deferred judgments, and whether a defendant’s felony drug case was handled in drug court or in a district court, but finding no difference in guilty plea charge reductions; interestingly Black people had their cases dismissed more often which raises questions about the police decision to arrest); KUTATELADZE ET AL., *supra* note 1, at 3 (finding Black and Latino people charged with misdemeanor drug offenses were more likely to have their cases dismissed, Black and Latino people with misdemeanor drug offenses were more likely to be detained at arraignment, and Black and Latino people with drug charges were more likely to receive more punitive plea offers and custodial sentences).

⁹⁹ See *supra* notes 2–5 and accompanying text.

¹⁰⁰ See NCSC REPORT, *supra* note 10, at I (noting that the report will “summarize[] what

report identifies three main types of interventions and explores their potential to reduce implicit bias.¹⁰¹ First, the report points to “intergroup contact” as a means of effectively reducing prejudice and discrimination.¹⁰² Intergroup contact means utilizing:

[e]ngagement activities that include the following features [in order to] have the greatest impact: 1) different groups are working toward a common goal, 2) the groups have equal status in the activity, 3) the activity allows individuals to get to know each other on an individual basis, and 4) the activity receives institutional support or support from the relevant authority figures.¹⁰³

Although the report does not specify a possible application in the legal profession, one can logically deduce that having diverse teams to make decisions within an organization will reduce the effects of implicit bias.¹⁰⁴

The NCSC report identifies a second technique as an effective means for reducing the effects of implicit biases on decision-making: limiting individual discretion as much as possible to reduce unwarranted disparities.¹⁰⁵ The report warns that “when individuals make decisions under conditions of limited structure, ambiguous decision-making procedures, or subjective criteria, they are more likely to make decisions that manifest their biases.”¹⁰⁶ Thus, the report recommends “[e]mbedding structure in the decision-making process, specifying decision-making procedures as clearly as possible, and relying more extensively on criteria that can be measured objectively.”¹⁰⁷ As discussed above, similar recommendations regarding the use of guidelines and checklists have been made for the medical profession.¹⁰⁸

is currently known from research in the psychological and brain sciences, including implicit bias strategies generally found to be effective and ineffective”).

¹⁰¹ *Id.* at 18.

¹⁰² *Id.* at 13.

¹⁰³ *Id.* at 18.

¹⁰⁴ The NCSC report mentions a study of jury decision-making that “found that juries composed of [w]hite and Black jurors engaged in higher-quality deliberations and made more egalitarian verdict decisions than juries composed of only [w]hite members.” *Id.* at 17. *See generally* Samuel R. Sommers, *On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations*, 90 J. PERSONALITY & SOC. PSYCH. 597 (2006). The report does not include this study as one exemplifying “intergroup contact,” but service on a diverse jury might qualify as intergroup contact. NCSC REPORT, *supra* note 10, at 17.

¹⁰⁵ NCSC REPORT, *supra* note 10, at 14.

¹⁰⁶ *Id.*

¹⁰⁷ *Id.*; *see also* Kovera, *supra* note 98, at 1154 (arguing for interventions that “constrain the ability of decision makers to act on their racial biases,” such as using “specific procedures for [] evaluations”).

¹⁰⁸ *See supra* notes 84–88 and accompanying text.

Third, the NCSC report finds that “interventions that [attempt to] bypass or disrupt biased responding show more promise.”¹⁰⁹ Bypassing interventions include blinding procedures such as the double-blind procedure in clinical studies, as well as those used in other professions, which reduce the incidence of error in the decision-making process by preventing consideration of irrelevant factors.¹¹⁰ The report notes that the National Research Council has recommended blind procedures when police administer eyewitness identification lineups.¹¹¹

The report considers interventions that “attempt to *change* implicit associations in memory [as] not [being] consistently effective.”¹¹² In this category, the report includes programs that involve “evaluative conditioning” or “exposing research participants to positive, counter-stereotypical exemplars,”¹¹³ and these approaches would presumably be part of implicit bias training programs. The NCSC report, like the medical literature, has likewise found implicit training programs to produce inconsistent outcomes,¹¹⁴ which stands to reason as there are many types of such programs. In general, researchers have found only minimal, short-term reductions of implicit bias, with inconsistent effects on decision-making, sometimes improving decision-making while other times spurring backlash.¹¹⁵

Like the NCSC report, legal scholars have also called for the adaptation of implicit bias reduction techniques for various aspects of the criminal justice system.¹¹⁶ It is important to consider the effects of implicit bias on various points in the criminal justice system since researchers have found that racial bias has a “huge cumulative discriminatory effect [that] was evident when all decision points were considered in totality—yielding a [forty-two] percent increased chance of a severe sanction for [B]lack defendants, even after controlling for legal factors.”¹¹⁷ We have

¹⁰⁹ See NCSC REPORT, *supra* note 10, at 18.

¹¹⁰ *Id.* at 16; see also *supra* notes 47–51 and accompanying text; Stefanie K. Johnson & Jessica F. Kirk, *To Reduce Gender Bias, Anonymize Job Applications*, HARV. BUS. REV. (Mar. 5, 2020), <https://hbr.org/2020/03/research-to-reduce-gender-bias-anonymize-job-applications> [<https://perma.cc/5MAV-SGU3>] (noting the use of blind auditions by symphony orchestras and advocating for the use of blind job applications in the early stages of a recruitment process to reduce bias). Johnson and Kirk emphasize the purpose of blinding as a form of quality control: “When you have evidence that gender, race, age, or other differences are affecting your selection process, despite their not being relevant selection criteria, you have *error* in your process.” *Id.*

¹¹¹ NCSC REPORT, *supra* note 10, at 16.

¹¹² See *id.* at 18; see also Kovera, *supra* note 98, at 1153 (finding “interventions designed to eliminate implicit bias or train decision makers to avoid it” as being “unlikely to eliminate racial disparities”).

¹¹³ NCSC REPORT, *supra* note 10, at 15.

¹¹⁴ *Id.* at 18; see also *supra* notes 78–79 and accompanying text.

¹¹⁵ NCSC REPORT, *supra* note 10, at 15–16.

¹¹⁶ See *infra* notes 118–25 and accompanying text.

¹¹⁷ See Christopher Robertson et al., *Race and Class: A Randomized Experiment with Prosecutors*, 16 J. EMPIRICAL LEGAL STUD. 807, 813 (2019) (addressing a meta-analysis

previously reported on the unique blind quality testing used by the City of Houston's crime laboratory, the Houston Forensic Science Center, and urged other laboratories to follow suit.¹¹⁸

Dan Simon has proposed blinding forensic death examiners from contextual information that is beyond their area of expertise, while allowing them to receive the types of background information within their expertise as a means of reducing context bias.¹¹⁹ Another scholar has proposed that judges use blinding techniques for case files to reduce their own implicit biases regarding individuals' "social category identity" (e.g., race), or that they use at least "temporary cloaking" (what scientists might call "sequential unmasking"¹²⁰) to remove information in the first stage of a two-stage process, and use checklists and rubrics to cabin their own discretion so as to make more uniform decisions in like situations.¹²¹ Another "double-blind" proposal suggests that defendants be kept out of view of the judge and jury during criminal trials, except when the defendant testifies.¹²²

Some scholars have recognized the importance of the "intergroup contact" discussed above in reducing law enforcement and prosecutorial bias, calling, for example, for increased diversity within prosecutor's offices.¹²³ Others have identified the prosecutorial charging decision as a pivotal decision point in the criminal justice and one where blinding case files to the race of a defendant could reduce the effects of implicit bias.¹²⁴ One previous laboratory study used blinding as an approach to

from sixty-five populous counties which found an effect of race at two of eight key decision points: likelihood of incarceration and length of sentence).

¹¹⁸ See Sandra Guerra Thompson & Nicole Bremner Cásarez, *Solving Daubert's Dilemma for the Forensic Sciences Through Blind Testing*, 57 HOUS. L. REV. 617, 618 (2020); THOMPSON, *supra* note 51, at 141–43 (calling for the use of blind procedures and sequential unmasking in forensic testing); see also *supra* note 49 and accompanying text (on sequential unmasking in medical research).

¹¹⁹ See Dan Simon, *Minimizing Error and Bias in Death Investigations*, 49 SETON HALL L. REV. 255, 255–56 (2019).

¹²⁰ See *supra* notes 49, 110 and accompanying text.

¹²¹ See Jerry Kang, *What Judges Can Do About Implicit Bias*, 57 CT. REV. 78, 83–86 (2021).

¹²² See Stanley P. Williams, Jr., *Double-Blind Justice: A Scientific Solution to Criminal Bias in the Courtroom*, 6 IND. J.L. & SOC. EQUAL. 48, 68–69 (2018).

¹²³ See, e.g., Melanie D. Wilson, *The Common Prosecutor*, 53 LOY. U. CHI. L.J. 325, 344–53 (2022); see also *supra* notes 100–04 and accompanying text (discussing the NCSC report's discussion of "intergroup contact" as a strategy for reducing implicit bias).

¹²⁴ See Kovera, *supra* note 98, at 1145 (noting studies that showed racial disparities in prosecutorial charging of juveniles and adults which had the effect of depriving them of more lenient punishment options); see Sunita Sah et al., *Blinding Prosecutors to Defendants' Race: A Policy Proposal to Reduce Unconscious Bias in the Criminal Justice System*, 1 BEHAV. SCI. & POL'Y 69, 69–76 (2015); Shima Baughman et al., *One Step Toward Making Criminal Justice Less Biased*, CONVERSATION (Oct. 13, 2016, 9:55 PM), <https://theconversation.com/one-step-toward-making-criminal-justice-less-biased-65022> [<https://perma.cc/VX7A-XAEX>]; Christina Couch, *To Reduce Bias, What if Prosecutors Didn't Know the Race of the People*

eliminate unwarranted race and class disparities in prosecutorial charging.¹²⁵ However, until recently, field studies in which prosecutors actually use race-blind case files have not been implemented.¹²⁶ In the section that follows, we report on two such studies.

B. Practical Applications of Bias Reduction Strategies in Criminal Justice: Race-Blind Prosecutorial Charging

Not until recently has a feasible and effective method existed to keep race-related information away from prosecutors until after they make their initial charging decisions.¹²⁷ In 2019, the San Francisco District Attorney's office (SFDAO) became the first in the nation to implement artificial intelligence blinding technology as a pilot program in an effort to eliminate implicit racial bias.¹²⁸ Also in California, the Yolo County District Attorney's Office (YCDAO) adopted the same blinding technology in May 2021, testing it by first applying it to cases referred to it by the Davis Police Department, and then expanding it in July 2021 to cases from the West Sacramento Police Department.¹²⁹ The YCDAO plans to extend race-blind charging to cases sent to it from four additional local police departments.¹³⁰

The technology, which was developed by the Stanford Computational Policy Lab at no cost to the district attorneys' offices,¹³¹ uses pattern-based and named entity recognition¹³² to automatically redact various categories of information contained

They Were Prosecuting?, FAST CO. (May 18, 2016), <https://www.fastcompany.com/3059872/to-reduce-bias-what-if-prosecutors-didnt-know-the-race-of-the-people-they-were-prosecuting> [<https://perma.cc/39JR-77J9>].

¹²⁵ See Robertson et al., *supra* note 117, at 818–22 (describing the methodology of using blinded case vignettes and asking participating prosecutors about how they would charge the individuals involved; the prosecutors had accepted an email invitation to participate in a study of prosecutor decision-making).

¹²⁶ See Sean Hollister, *San Francisco Says It Will Use AI to Reduce Bias When Charging People with Crimes*, VERGE (June 12, 2019, 2:57 PM), <https://www.theverge.com/2019/6/12/18663093/ai-sf-district-attorney-police-bias-race-charge-crime> [<https://perma.cc/X8PK-FXER>].

¹²⁷ Before 2019, the San Francisco District Attorney's Office tried to keep prosecutors from learning race-related information from police incident reports by manually removing the first two pages of the report. However, prosecutors had full access to the remainder of the police report. *See id.*

¹²⁸ *Id.*

¹²⁹ See Lauren Keene, *Yolo DA Unveils Race-Blind Charging Program*, DAVIS ENTER. (Sept. 10, 2021), <https://www.davisenterprise.com/news/local/yolo-da-unveils-race-blind-charging-program/> [<https://perma.cc/7PC7-E54G>].

¹³⁰ See Sam Stanton, *Northern California DA Using Software to Make 'Race Blind' Decisions on Filing Charges*, SACRAMENTO BEE (Sept. 9, 2021, 5:00 AM), <https://www.sacbee.com/news/local/article253708533.html> [<https://perma.cc/UVP2-CQN9>].

¹³¹ *See id.*; Hollister, *supra* note 126.

¹³² "Named entity recognition" is a natural language processing technique through which

in police incident reports that could reveal or suggest an individual's race.¹³³ Specifically, the Stanford masking algorithm identifies and conceals five categories of information contained in felony police incident reports: (1) explicit references to race; (2) physical descriptions, including hair and eye color; (3) names and nicknames; (4) location information, including street addresses and names of neighborhoods; and (5) identifying information about police officers, because a prosecutor who knows where an officer is stationed may be able to infer the neighborhood where the incident occurred.¹³⁴

Rather than simply deleting this data, or obscuring it with black-box redactions, the blinding tool indicates the type of information that has been removed from the report in a generic manner.¹³⁵ For example, if the original police report mentioned "a Black male," the redacted report would instead reference "a [race] male."¹³⁶ Individuals are given racially neutral labels that describe their role in a case, such as "Victim 1" or "Officer 1," and are highlighted in a particular color that remains constant throughout the redacted report.¹³⁷ This helps retain the report's narrative structure, and ensures that prosecutors can still read and comprehend the report easily.¹³⁸

words that identify people, job titles, objects, places, etc. are recognized and extracted from text, and classified into predefined categories. *See, e.g.*, Arun Jagota, *Named Entity Recognition in NLP*, TOWARDS DATA SCI. (July 9, 2020), <https://towardsdatascience.com/named-entity-recognition-in-nlp-be09139fa7b8> [<https://perma.cc/5CYP-RXPB>]. Pattern-based recognition is used to identify and extract information from text using structural patterns—for example, ZIP Codes and telephone numbers—rather than by words. *Id.*

¹³³ Maria Dinzeo, *San Francisco DA Turns to AI to Tame Racial Bias*, COURTHOUSE NEWS SERV. (June 12, 2019), <https://www.courthousenews.com/san-francisco-da-turns-to-ai-to-tame-racial-bias/> [<https://perma.cc/X9H5-8XGG>].

¹³⁴ Alex Chohlas-Wood et al., *Blind Justice: Algorithmically Masking Race in Charging Decisions*, in PROCEEDINGS OF THE 2021 AAAI-ACM CONFERENCE ON AI, ETHICS, & SOCIETY 35, 36 (2021), <https://dl.acm.org/doi/abs/10.1145/3461702.3462524> [<https://perma.cc/BBA8-J3RK>].

The tool masks only information that could pertain to race, as opposed to other demographic data, from police incident reports. *Id.* at 38 (explaining that some demographic information can have a "direct bearing on whether a prosecutor decides to charge a case" and stating that, for example, "a victim's gender may inform whether a physical altercation was mutually provoked or more likely one-sided").

Although the Stanford researchers did not name the jurisdiction where their algorithm initially had been implemented, the San Francisco District Attorney held a widely reported press conference on June 12, 2019, to announce plans to launch the program. *See, e.g.*, Dinzeo, *supra* note 133. One press report noted another reason that the tool removes details about police officers was to eliminate any bias prosecutors may have either for or against reports written by certain officers. Hollister, *supra* note 126.

¹³⁵ Chohlas-Wood et al., *supra* note 134, at 38.

¹³⁶ *Id.* Color descriptions in police reports that do not relate to race, e.g., a reference to a black jacket or a white car, are not removed. *Id.*

¹³⁷ *Id.*

¹³⁸ *Id.*

However, the tool eliminates suspects' criminal histories from the incident reports as part of the redaction.¹³⁹

To implement race-blind charging using the Stanford algorithm, district attorneys employ a two-step case assessment procedure.¹⁴⁰ First, prosecutors are instructed to make and document their initial charging decisions based solely on the redacted incident reports, and before speaking with any police officers.¹⁴¹ In Yolo County, this means that prosecutors choose between four options regarding how likely they are to bring charges against the suspect: very likely, likely, unlikely, or very unlikely.¹⁴² As part of this initial review, prosecutors are also asked whether the algorithm did a "good" or "bad" job scrubbing racial cues from the police report.¹⁴³ The researchers use this feedback to modify the algorithm if it redacts too much or too little information.¹⁴⁴

After recording their initial charging decisions, prosecutors are furnished with full, unredacted police reports, including any photographic or video evidence and the suspect's criminal history, and are asked for final charging decisions in the case.¹⁴⁵ If their final decision differs from their decision on blind review, they are required to explain and document their reasons for the change.¹⁴⁶ According to George Gascon, the former San Francisco District Attorney who initiated the SFDAO pilot program, the explanation requirement encourages prosecutors to stand by their first, color-blind determinations.¹⁴⁷

The YCDAO predicted that about seventy to seventy-five percent of misdemeanor and felony cases submitted to it would be subject to the race-blind charging program.¹⁴⁸ Homicide, domestic violence and sexual assault cases are excluded from the race blind review, because YCDAO prosecutors rely more heavily on rap sheets,

¹³⁹ See *Yolo County Launches Race-Blind Charging Program to Remove Biases from Criminal Justice System*, CBS SACRAMENTO (Sept. 16, 2021, 11:18 AM) [hereinafter *Yolo County*], <https://sacramento.cbslocal.com/2021/09/16/yolo-county-launches-race-blind-charging-program-to-remove-biases-from-criminal-justice-system/> [<https://perma.cc/8ZYF-BTF5>].

¹⁴⁰ Chohlas-Wood et al., *supra* note 134, at 38.

¹⁴¹ *Id.* at 39. At this stage, prosecutors choose between four options: "charge," "probably charge," "probably discharge" and "discharge," and provide a brief explanation for their choice. *Id.*

¹⁴² See Keene, *supra* note 129.

¹⁴³ *Id.* If information that is supposed to be removed nevertheless survives the redaction process, Yolo County prosecutors charge that case in the traditional manner. See Stanton, *supra* note 130.

¹⁴⁴ Kenrick Cai, *San Francisco Will Use AI to Thwart Racial Bias When Charging Suspects*, FORBES (June 12, 2019, 9:48 PM), <https://www.forbes.com/sites/kenrickcai/2019/06/12/san-francisco-will-use-ai-to-thwart-racial-bias-when-charging-suspects/?sh=2d9620712bfd> [<https://perma.cc/672V-WQY8>].

¹⁴⁵ Chohlas-Wood et al., *supra* note 134, at 39.

¹⁴⁶ *Id.*

¹⁴⁷ See Cai, *supra* note 144.

¹⁴⁸ Keene, *supra* note 129.

DNA profiles, crime-scene photographs and victim interviews to make charging decisions in those types of cases.¹⁴⁹ These cases, as well as officer-involved shootings and police use-of-force cases, were also excluded from the SFDAO's pilot blind review program, in part because, with respect to the latter, prosecutors needed to be aware of an individual's status as a law enforcement official.¹⁵⁰

The SFDAO's blind-review pilot program operated from August 2019 through March 2020, when it was discontinued at the start of the COVID-19 pandemic.¹⁵¹ During that time, SFDAO prosecutors conducted more than four hundred preliminary blind reviews, which the Stanford researchers statistically analyzed with respect to (1) how well the algorithm concealed a suspect's race in police incident reports; and (2) how the blind review process affected prosecutors' charging decisions.¹⁵² To determine the quality of the redactions, they engaged an examiner to review redacted police reports and predict whether a particular incident involved a Black suspect.¹⁵³ They also trained several machine-learning models¹⁵⁴ to do the same thing.¹⁵⁵ Based on those results, the researchers concluded that algorithmic redaction effectively suppressed racial cues present in the reports.¹⁵⁶

To evaluate the two-step case assessment procedure on prosecutor's charging decisions, the researchers compared charging decisions in cases that included the preliminary blind review to those that did not, and observed only small differences between the two.¹⁵⁷ In other words, charging rates were comparable whether San Francisco prosecutors reviewed redacted or unredacted case files.¹⁵⁸ Furthermore, while prosecutors initially recommended pressing charges in sixty-eight percent of cases following blind review, that charging rate dropped to fifty-seven percent after prosecutors were allowed to examine the complete, unredacted case file.¹⁵⁹ The researchers advanced three possible reasons for this reduction: first, that by removing

¹⁴⁹ *Id.*

¹⁵⁰ James Queally, *San Francisco D.A. Unveils Program Aimed at Removing Implicit Bias from Prosecutions*, L.A. TIMES (June 12, 2019, 11:00 AM), <https://www.latimes.com/local/lanow/la-me-san-francisco-da-prosecutions-implicit-bias-software-20190612-story.html> [<https://perma.cc/UC25-UEAP>].

¹⁵¹ Telephone Interview with Alex Chohlas-Wood, Executive Director, Stanford Computational Policy Lab (Sept. 27, 2021) (notes on file with authors) [hereinafter Chohlas-Wood Interview].

¹⁵² Chohlas-Wood et al., *supra* note 134, at 40.

¹⁵³ *Id.* at 39.

¹⁵⁴ A machine-learning model is a computer program that has been trained to recognize certain types of patterns in data. See *What Is a Machine Learning Model?*, MICROSOFT (Dec. 30, 2021), <https://docs.microsoft.com/en-us/windows/ai/windows-ml/what-is-a-machine-learning-model> [<https://perma.cc/4NR4-ZQ9R>].

¹⁵⁵ Chohlas-Wood et al., *supra* note 134, at 39.

¹⁵⁶ *Id.* at 40.

¹⁵⁷ *Id.* at 41.

¹⁵⁸ *Id.* at 36.

¹⁵⁹ *Id.* at 42.

personal information from the incident report, the redaction process might have caused prosecutors to respond more harshly than they otherwise would; second, that prosecutors might have wrongly assumed the existence of incriminating evidence that they then failed to find in the full case file, leading them to ultimately drop the charges; and third, without knowing all the evidence, prosecutors may be more likely to default to conservative behavior and charge suspects.¹⁶⁰

While the program's overall lack of effect on the SFDAO's overall charging decisions might be surprising to some, earlier studies of the SFDAO had reported little historic evidence of racially disparate treatment well before adoption of the blinding algorithm.¹⁶¹ Accordingly, the researchers concluded that their results simply provided more evidence that "race does not substantially impact charging decisions" made by the SFDAO.¹⁶² This may well not be true in other jurisdictions, as other studies have suggested,¹⁶³ meaning that prosecutorial blinding could play a more significant role in reducing racially disparate treatment in charging decisions elsewhere. Similarly, the Yolo County District Attorney has said that although he does not expect that the race-blind charging program will significantly affect charging practices in his office, the program's implementation will help counter public distrust of the criminal justice system.¹⁶⁴

Whether race-blind charging can make a significant difference in other jurisdictions may be answered relatively soon, given that in April 2022, the Yolo County District Attorney co-sponsored a bill that would require all district attorneys' offices in California to implement race-blind charging systems by January 2025.¹⁶⁵ Regardless

¹⁶⁰ *Id.* According to the Stanford Computational Policy Lab's executive director, Alex Chohlas-Wood, written explanations provided by SFDAO prosecutors regarding changes in charging decisions failed to show any strong patterns. Chohlas-Wood Interview, *supra* note 151.

¹⁶¹ Chohlas-Wood et al., *supra* note 134, at 37; John MacDonald & Steven Raphael, *Effect of Scaling Back Punishment on Racial and Ethnic Disparities in Criminal Case Outcomes*, 19 CRIMINOLOGY & PUB. POL'Y 1139, 1151 (2020); see also EMILY OWENS ET AL., EXAMINING RACIAL DISPARITIES IN CRIMINAL CASE OUTCOMES AMONG INDIGENT DEFENDANTS IN SAN FRANCISCO 20 (2017), <https://www.law.upenn.edu/live/files/6793-examining-racial-disparities-may-2017-full> [<https://perma.cc/Y62D-5YF2>] (concluding, based on 10,753 complete case records from 2011–2014 from the San Francisco Public Defender's Office, that "[no] evidence [exists] that district attorneys file more or fewer charges against Black or Latinx defendants [than they file against whites].").

¹⁶² Chohlas-Wood et al., *supra* note 134, at 36.

¹⁶³ See, e.g., M. Marit Rehavi & Sonja B. Starr, *Racial Disparity in Federal Criminal Sentences*, 122 J. POL. ECON. 1320, 1350 (2014) (concluding that differences in prosecutors' initial charging decisions was the primary cause of sentencing disparities between Black and white defendants in the federal system). *But see* Robertson et al., *supra* note 117, at 845 (finding no evidence of racially disparate charging decisions by prosecutors).

¹⁶⁴ See Stanton, *supra* note 130; *Yolo County*, *supra* note 139 (noting that the YCDAO decided to adopt blind charging after meeting with a local citizen's group that expressed misgivings regarding the fairness of the criminal process to people of color).

¹⁶⁵ See Press Release, Cnty. of Yolo, Off. of the Dist. Att'y, Yolo DA Sponsors Criminal

of whether that bill ultimately becomes California law, the Stanford Computational Policy Lab hopes to implement its algorithm next in counties that have a history of racially disparate charging.¹⁶⁶ Assuming that the technology can be effective in reducing racially disparate charging, the Stanford researchers could make the algorithm freely available as open source software, and have considered embedding the algorithm in commercially available case management software already utilized by district attorneys' offices.¹⁶⁷ Although many other district attorneys' offices have contacted the researchers to express interest in implementing the free race-blind charging program, lack of technology has been the biggest obstacle to adoption.¹⁶⁸ To utilize this technology in casework, district attorneys' offices must first have access to a digital feed of incident report data from the police department.¹⁶⁹ For example, before implementing the race-blind charging pilot program, the SFDAO first had to convert from a paper-based to a digitized system.¹⁷⁰

The Stanford algorithm is not the only technological solution that has been proposed to assist district attorneys' offices with eliminating racial bias in charging decisions. Researchers at the Princeton Project in Computational Law have proposed that district attorneys' offices review suggestions from three artificial-intelligence-based computational models created from historic data when making charging or sentencing decisions.¹⁷¹ The first of these models, the "predictive" model, reflects how the suspect would have been treated historically by that district attorney's office.¹⁷² The second, the "race-neutral" model, shows how the suspect would have been charged historically without specific reference to race, but including other demographic data such as neighborhood that can serve as a proxy for race.¹⁷³ Finally, the third, the "suggestive" model, informs the prosecutor how the defendant would have been charged historically had he or she been white.¹⁷⁴ The "suggestive" model was developed as a response to a justifiable criticism of using artificial intelligence as

Reform Bill (Apr. 20, 2022), <https://yoloda.org/wp-content/uploads/2022/04/PRESS-RELEASE-RACE-BLIND-CHARGING-AB-2418-PASSES-ASSEMBLY-PUBLIC-SAFETY-COMMITTEE-1.pdf> [<https://perma.cc/75B8-HDKA>].

¹⁶⁶ See Gerardo Zavala, *Yolo County DA Announces Launch of Race Blind Program for Filing Charges*, DAILY DEMOCRAT (Sept. 9, 2021, 4:45 PM), <https://www.dailydemocrat.com/2021/09/09/yolo-county-da-announces-launch-of-race-blind-program-for-filing-charges/> [<https://perma.cc/YF6B-GT5F>].

¹⁶⁷ Chohlas-Wood Interview, *supra* note 151.

¹⁶⁸ *Id.*

¹⁶⁹ *Id.*

¹⁷⁰ Chohlas-Wood et al., *supra* note 134, at 36.

¹⁷¹ JOSEPH J. AVERY ET AL., USING ARTIFICIAL INTELLIGENCE TO IMPROVE THE FAIRNESS AND EQUITY OF GOVERNMENT DECISION MAKING 40 (2020), https://napawash.org/uploads/Using_AI_to_Improve_the_Fairness_and_Equity_of_Government_Decision_Making.pdf [<https://perma.cc/9P4A-HGQ6>].

¹⁷² *Id.*

¹⁷³ *Id.*

¹⁷⁴ *Id.*

a bias reduction tool¹⁷⁵: that, at least in some instances, algorithms based on historic data merely replicate and reinforce existing racial disparities.¹⁷⁶

While the Princeton proposal could be adopted in paper-based, rather than digitized, district attorneys' offices, it nevertheless presents other significant hurdles to implementation. Many district attorneys' offices lack the financial resources and/or technical know-how to design, build and apply such a complex artificial intelligence system that, of necessity, would be unique to each office.¹⁷⁷ Furthermore, the three proposed computational models cannot be created unless the district attorney's office has collected accessible, detailed historic charging and/or sentencing data.¹⁷⁸

The fear that the use of artificial intelligence will exacerbate historical bias is not an issue, however, with the Stanford blinding algorithm, which merely redacts information moving forward.¹⁷⁹ It is not predictive and does not use historic data for any purpose.¹⁸⁰ As the researchers themselves have explained, using the blinding tool is equivalent to hiring human beings to manually redact police incident reports as they are submitted, only without the accompanying overhead expense and time commitment.¹⁸¹

Of course, as the Stanford researchers have emphasized, a prosecutor's initial charging decision is only one of many stages in criminal justice process where implicit bias can result in racially disparate impacts.¹⁸² For example, in San Francisco, a 2017 report by the Quattrone Center concluded that more serious charges imposed against people of color by police at the initial booking stage had a greater impact with regard to racial disparity than charges brought by the SFDAO.¹⁸³ In an attempt to detect and address implicit bias in cases presented to the SFDAO by law enforcement, the District Attorney in 2020 responded with the NCSC's recommended technique of limiting individual prosecutorial discretion: he implemented a policy against filing possession of contraband charges in cases where an officer uses a minor traffic infraction to stop and search a motorist.¹⁸⁴ These "pretextual" traffic

¹⁷⁵ *Id.* at 41.

¹⁷⁶ See, e.g., Vincent Southerland, *With AI and Criminal Justice, the Devil Is in the Data*, ACLU (Apr. 9, 2018, 11:00 AM), <https://www.aclu.org/issues/privacy-technology/surveillance-technologies/ai-and-criminal-justice-devil-data> [<https://perma.cc/PU9R-TXNR>].

¹⁷⁷ See *Joint Panel Session Between the Standing Panel on Social Equity in Governance and the Standing Panel on Technology Leadership: Can AI Be Used to Increase Fairness and Equity in Government Decisions?*, NAT'L ACAD. OF PUB. ADMIN. (Dec. 28, 2020), <https://napawash.org/standing-panel-blog/can-ai-be-used-to-increase-fairness-and-equity-in-government-decisions> [<https://perma.cc/ZG67-XT9H>].

¹⁷⁸ See *id.* (noting that "[m]ost localities don't know the data they have").

¹⁷⁹ Chohlas-Wood et al., *supra* note 134, at 37.

¹⁸⁰ *Id.*

¹⁸¹ *Id.*

¹⁸² *Id.* at 44.

¹⁸³ OWENS ET AL., *supra* note 161, at 24. In one press account, the former San Francisco District Attorney who implemented the pilot program suggested that it could also be a way for the police "to see their own work." See Dinzeo, *supra* note 133.

¹⁸⁴ See Press Release, S.F. Dist. Att'y, District Attorney Boudin Pioneers First in the Nation

stops were found to be used by police most often against Black drivers in San Francisco, who are three times more likely to be searched than white drivers.¹⁸⁵ Additionally, to do a better job identifying those cases, the SFDAO in 2021 applied for and received grant funding and technical assistance from Berkeley's The People Lab to (1) conduct research to identify the characteristics of pretextual stops; and (2) to develop an AI algorithm to flag arrest reports that reflect those characteristics.¹⁸⁶ In this way, the SFDAO hopes to improve its ability to "catch" all arrests based on pretextual stops and thereby reduce implicit bias at another critical juncture of the criminal justice process.¹⁸⁷

CONCLUSION

Obvious racial disparities in the criminal justice system have reduced public trust in law enforcement and the courts, as well as sparked urgent calls for reform. We can agree that racial bias, like socioeconomic bias, gender bias, ageism, attractiveness bias, or any other type of bias, should almost never play a role in criminal justice outcomes.¹⁸⁸ This Article has suggested that implicit racial bias in criminal justice be addressed not as a moral failing of the system's participants, but rather in a scientific manner—as an unavoidable cause of error that must be methodically minimized to ensure valid, trustworthy outcomes.

Medical researchers and other scientists continually develop, study and test strategies to address the implicit biases that threaten the validity of their scientific work. They favor a layered, reinforcing approach, recognizing that bias can emerge at many points in any complex process that depends on discretionary decision-making. They recognize that no "magic bullet" will provide a one-shot fix. The parallels with, and the lessons for, the criminal justice process are clear. We believe that by emulating the scientific approach to bias reduction, we will be better equipped to devise effective bias minimization strategies to improve criminal justice.

Policy Directives (Feb. 28, 2020), <https://www.sfdistrictattorney.org/press-release/da-boudin-pioneers-first-in-nation-policy-directives/> [<https://perma.cc/QUP2-9ZGG>].

¹⁸⁵ *Id.*

¹⁸⁶ See E-mail from Nicole Casarez, Emeritus Professor, Univ. St. Thomas, to Mikaela Rabinowitz, Dir. of Data, Rsch., & Analytics, S.F. Dist. Att'y's Off. (Feb. 22, 2022, 2:12 PM) (on file with authors).

¹⁸⁷ *Id.* In May 2022, the San Francisco Police Commission announced that it was considering a proposal to limit a police officer's ability to pull over motorists for low-level traffic offenses, in its own effort to reduce the number of pretextual stops. See Megan Cassidy, *Why S.F. Might Be About to Prohibit Police from Making Low-Level Traffic Stops*, S.F. CHRON. (May 11, 2022, 6:36 PM), <https://www.sfchronicle.com/sf/article/Why-S-F-might-be-about-to-prohibit-police-from-17166395.php> [<https://perma.cc/6DVY-J8T2>].

¹⁸⁸ An exception to this precept might be made for an extremely small number of cases, such as "hate crimes," for example, where race would be a salient consideration.

Fortunately, researchers have begun to examine and identify those intervention techniques that hold the most promise for reducing implicit bias at various key points in the criminal justice process. As described in this Article, a few forward-thinking prosecutors have implemented discretion-limiting policies as well as advanced blinding technologies in an attempt to reduce racial disparities in their jurisdictions. Although the results so far have either been mixed or are still unknown, we applaud their willingness to collaborate with university research labs and to foster a research culture in their offices. Only by employing bias reduction techniques, measuring their effects and sharing those results—in other words, applying the trial-and-error process implicit in the scientific method—will criminal justice reformers succeed in significantly reducing implicit bias in the criminal justice system.