RACIAL BIAS IN POST-ARREST AND PRETRIAL DECISION MAKING: THE PROBLEM AND A SOLUTION

Joseph J. Avery & Joel Cooper***

There is significant racial disparity in U.S. incarceration rates, with African Americans vastly overrepresented. Given that almost all criminal cases resolve through plea bargaining, a process that takes place between the poles of arrest and trial, if progress concerning racial disparity in incarceration rates is to be made, greater understanding of the post-arrest and pretrial (PAPT) period is necessary. What do we know about racial bias in PAPT decision making, and what can the psychology of prejudice teach us about it?

In this Article, we map the factors that are liable to lead individuals to disparate treatment decisions, and we apply this thinking to the primary institutional PAPT actors: prosecutors, criminal defense attorneys, and judges. We also consider defendants' attitudes towards the justice system and their perceptions of discrimination, as well as what such cognitions mean for their plea bargaining decisions.

Finally, while a number of interventions have been tried and even more have been proposed, very little progress in racial disparity in incarceration rates has been made. We endorse a novel way forward. Prosecutors should be provided with race-neutral baselines, anchors to their decision making as a case progresses, and prosecutors' offices should organize their institutional data and allow either internal or external parties to analyze it and produce regular and open reports. In

---

* Woodrow Wilson Scholars Fellow, Princeton University; Columbia Law School, J.D.; Princeton University, PhD Candidate; Princeton University, M.A.; New York University, B.A. This research was partially supported by a grant from the Horowitz Foundation for Social Policy, and Mr. Avery was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program. We are grateful to all those who took the time to provide comments, critiques, and suggestions on this research, including J. Nicole Shelton, the members of Princeton's Stigma and Social Perception Lab, and the members of the Princeton Project in Computational Law. In addition, we are grateful for the support provided by the *Cornell Journal of Law & Public Policy*, especially Ferdinand Suba, Jr., and Andrew Steele.

** Professor of Psychology, Princeton University; Duke University, Ph.D.; City College of the City University of New York, B.A.

addition to resulting in greater fairness in PAPT decision making, this process will increase trust in the U.S. criminal justice system.

INTRODUCTION

In Franz Kafka's *The Trial*,[1] the horror in the story comes from the inscrutable process set in motion after a seemingly innocent man is arrested. The opening sentence describes Joseph K.'s arrest and the concluding sentence his execution.[2] What comes between is the substance of the novel: the crime of which Joseph K. is accused is never revealed, and the authorities who pronounce his guilt are unseen, their reasoning unclear.[3] Joseph K. encounters an impenetrable justice system, its true motivations and workings beyond his reach. In her 2016 book, Nicole Gonzalez Van Cleve picks up this same thread. She provides a journalistic exploration of injustice in post-arrest and pretrial processes in an Illi-

[1] *See* FRANZ KAFKA, THE TRIAL (Mike Mitchell, trans., Oxford University Press 2009) (1937).

[2] *Id.* at 5, 165.

[3] *See id.*

nois county.[4] In the 1980s and 1990s, prosecutors in Cook County played a game they called "The Two-Ton Contest." The goal of the game was to imprison 4,000 pounds of African American individuals as quickly as possible.[5]

More than 2 million people are incarcerated in the United States,[6] and a disproportionate number of these individuals are African Americans. In 2015, African Americans made up 13% of the U.S. population[7] and 37% of the incarcerated population.[8] One in 3 African American males will go to prison during their lifetimes, while the same is true for only 1 in 17 white males.[9] In a 2015 speech, then-President Obama cited the disproportionate number of minority men in prison in concluding that the criminal justice system was broken.[10] Today, there is widespread, bipartisan agreement with Obama's conclusion.[11]

The stakes of improper or unjust convictions are high—the American Bar Association identified more than 38,000 punitive provisions that apply to people convicted of crimes, including loss of public housing, welfare assistance, and occupational licenses,[12] as well as loss of financial aid for higher education and even social security benefits.[13] Felons may lose the right to vote.[14] Incarceration itself poses significant health

---

4 NICOLE GONZALEZ VAN CLEVE, CROOK COUNTY: RACISM AND INJUSTICE IN AMERICA'S LARGEST CRIMINAL COURT 3 (2016).

5 *Id.* at 54.

6 *Overview: United States of America*, WORLD PRISON BRIEF (last updated Dec. 31, 2016), http://www.prisonstudies.org/country/united-states-america.

7 *Annual Estimates of the Resident Population by Sex, Single Year of Age, Race Alone or in Combination, and Hispanic Origin for the United States: April 1, 2010 to July 1, 2015*, U.S. CENSUS BUREAU: AM. FACTFINDER (June 2016), https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=PEP_2015_PEPALL5N&prodType=table.

8 E. ANN CARSON, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 248955, PRISONERS IN 2014 15 (2015), https://www.bjs.gov/content/pub/pdf/p14.pdf.

9 THOMAS P. BONCZAR, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 197976, PREVALENCE OF IMPRISONMENT IN THE U.S. POPULATION, 1974–2001 1 (2003), https://www.bjs.gov/content/pub/pdf/piusp01.pdf; *see also* E. ANN CARSON, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 247282, PRISONERS IN 2013 8 (2014), https://www.bjs.gov/content/pub/pdf/p13.pdf (showing prison populations in 2013 by race).

10 Peter Baker, *Obama Calls for Effort to Fix a 'Broken System' of Criminal Justice*, N.Y. TIMES (July 14, 2015), https://www.nytimes.com/2015/07/15/us/politics/obama-calls-for-effort-to-fix-a-broken-system-of-criminal-justice.html.

11 Kathryne M. Young & Joan Petersilia, *Keeping Track: Surveillance, Control, and the Expansion of the Carceral State*, 129 HARV. L. REV. 1318, 1318 (2016) (book review).

12 Alfred Blumstein & Kiminori Nakamura, *Paying a Price, Long After the Crime*, N.Y. TIMES (Jan. 9, 2012), https://www.nytimes.com/2012/01/10/opinion/paying-a-price-long-after-the-crime.html.

13 JAMES B. JACOBS, THE ETERNAL CRIMINAL RECORD 257–59 (2015).

14 *Id.* at 249.

risks.[15] Inmates routinely face sexual abuse,[16] and they often fail to receive proper mental health care and treatment.[17] After release, something as essential and as seemingly quotidian as finding employment becomes difficult if not impossible.[18]

Given the salience of unjust arrests and unjust trials,[19] these stages in the criminal justice process have received significant attention from scholars and policy experts.[20] But racial bias that emerges between these two poles is arguably more important. Approximately 95-97% of convic-

---

15 *See* NATIONAL RESEARCH COUNCIL & INSTITUTE OF MEDICINE, HEALTH AND INCARCERATION: A WORKSHOP SUMMARY 7–8 (2013); *see also* Kenneth L. Appelbaum et al., *A National Survey of Self-Injurious Behavior in American Prisons*, 62 PSYCHIATRIC SERVS. 285, 285, 289 (2011); *see also* Craig Haney, *Mental Health Issues in Long-Term Solitary and "Supermax" Confinement*, 49 CRIME & DELINQ. 124, 125, 132 (2003); Ariel Ludwig et al., *Injury Surveillance in New York City Jails*, 102 AM. J. PUB. HEALTH 1108, 1108–09 (2012).

16 *See* ALLEN J. BECK, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 238640, PREA DATA COLLECTION ACTIVITIES, 2012 1–2 (2012), https://www.bjs.gov/content/pub/pdf/pdca12.pdf; *see also* ALLEN J. BECK & PAIGE M. HARRISON, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 221946, SEXUAL VICTIMIZATION IN LOCAL JAILS REPORTED BY INMATES, 2007 (2008), https://www.bjs.gov/content/pub/pdf/svljri07.pdf; *see also* Nancy Wolff & Jing Shi, *Contextualization of Physical and Sexual Assault in Male Prisons: Incidents and Their Aftermath*, 15 J. CORRECTIONAL HEALTH CARE 58, 64 (2009).

17 *See* Carrie B. Oser et al., *Organizational-Level Correlates of the Provision of Detoxification Services and Medication-Based Treatments for Substance Abuse in Correctional Institutions*, 103 DRUG AND ALCOHOL DEPENDENCE S73, S74 (2009); *see also* Josiah D. Rich, Sarah E. Wakeman & Samuel L. Dickman, *Medicine and the Epidemic of Incarceration in the United States*, 364 NEW ENG. J. MED. 2081, 2081–82 (2011); *see also* Editorial Bd., *If Addiction Is a Disease, Why Is Relapsing a Crime?*, N.Y. TIMES (May 29, 2018), https://www.nytimes.com/2018/05/29/opinion/addiction-relapse-prosecutions.html.

18 *See* Young & Petersilia, *supra* note 11, at 1336–40.

19 *See generally* DEVERY S. ANDERSON, EMMETT TILL: THE MURDER THAT SHOCKED THE WORLD AND PROPELLED THE CIVIL RIGHTS MOVEMENT (2015); David C. Baldus et al., *Evidence of Racial Discrimination in the Use of the Death Penalty: A Story from Southwest Arkansas (1990-2005) with Special Reference to the Case of Death Row Inmate Frank Williams, Jr.*, 76 TENN. L. REV. 555 (2009); Sheri Lynn Johnson et al., *The Delaware Death Penalty: An Empirical Study*, 97 IOWA L. REV. 1925 (2012); James S. Liebman & Peter Clarke, *Minority Practice, Majority's Burden: The Death Penalty Today*, 9 OHIO ST. J. CRIM. L. 255 (2011); Glenn L. Pierce & Michael L. Radelet, *Death Sentencing in East Baton Rouge Parish, 1990-2008*, 71 LA. L. REV. 647 (2011); Adam Liptak, *Death Penalty Case Heard by Racist Juror Is Reopened by Supreme Court*, N.Y. TIMES (Jan. 8, 2018), https://www.nytimes.com/2018/01/08/us/politics/death-penalty-case-heard-by-racist-juror-is-reopened-by-supreme-court.html; Ashley Southall, *To Curb Bad Verdicts, Court Adds Lesson on Racial Bias for Juries*, N.Y. TIMES (Dec. 15, 2017), https://www.nytimes.com/2017/12/15/nyregion/to-curb-bad-verdicts-court-adds-lesson-on-racial-bias-for-juries.html.

20 Edith Greene et al., *Jurors and Juries: A Review of the Field*, *in* TAKING PSYCHOLOGY AND LAW INTO THE TWENTY-FIRST CENTURY 225 (James R. P. Ogloff ed., 2002); Justin D. Levinson et al., *Judging Implicit Bias: A National Empirical Study of Judicial Stereotypes*, 69 FLA. L. REV. 63, 65–66 (2017); Jill K. Swencionis & Phillip A. Goff, *The Psychological Science of Racial Bias and Policing*, 23 PSYCHOL., PUB. POL'Y, & L. 398, 398 (2017). *See generally* Dennis J. Devine et al., *Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups*, 7 PSYCHOL., PUB. POL'Y, & L. 622, 622 (2001).

tions are obtained by guilty plea.[21] In short, almost all criminal cases resolve in the nebulous period described in Kafka's fiction and in Van Cleve's journalism—the interstice between arrest and trial. The rules governing this period are unclear, and the primary authorities are largely immune from oversight. If one seeks to lessen racial disparities in incarceration rates, this is the period that needs to be better understood and better addressed.

Two questions should be asked. First, what is the role of racial bias in post-arrest and pretrial (PAPT) processes? Second, what should be done to correct for racial bias in PAPT decision making? As per the first question, there is a robust literature concerning bias and prejudice,[22] one that outlines cognitive processes that lead to prejudice,[23] situational factors that impact the presentation of bias,[24] the role played by individual differences,[25] and the ways in which social and public institutions create, justify, and perpetuate bias.[26] Moreover, as we will discuss, some researchers have even trained a lens on specific PAPT legal actors.

In this Article, we provide an overview of how racial bias in the PAPT period affects the outcomes of criminal cases. We will focus on four primary actors: prosecutors, criminal defense attorneys, judges, and defendants. Given that much of the research in law and the social sci-

---

21 ANN L. PASTORE & KATHLEEN MAGUIRE, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS 1, 443–50 (2003); U.S. DEP'T JUST. EXEC. OFFICE FOR U.S. ATTYS, U.S. ATTORNEY'S ANNUAL STATISTIC REPORT: FISCAL REPORT 2013 1, 9 (2014), https://www.justice.gov/sites/default/files/usao/legacy/2014/09/22/13statrpt.pdf; MATTHEW R. DUROSE & PATRICK A. LANGAN, U.S. DEP'T OF JUSTICE, BUREAU OF JUSTICE STATISTICS BULL., FELONY SENTENCES IN STATE COURTS, 2004, at 1, (July 2007), https://www.bjs.gov/content/pub/pdf/fssc04.pdf; MARK MOTIVANS, U.S. DEP'T OF JUSTICE, BUREAU OF JUSTICE STATISTICS, FEDERAL JUSTICE STATISTICS, 2012: STATISTICAL TABLES, at 17 (2015), https://www.bjs.gov/content/pub/pdf/fjs12st.pdf. Plea bargaining is a process wherein a defendant receives less than the maximum charge possible in exchange for an admission of guilt or something functionally equivalent to guilt. *See generally* THOMAS H. COHEN & BRIAN A. REAVES, U.S. DEP'T OF JUSTICE, BUREAU OF JUSTICE STATISTICS, FELONY DEFENDANTS IN LARGE URBAN COUNTIES, 2002, at i, iv. (2006), https://www.bjs.gov/content/pub/pdf/fdluc02.pdf. Cases that resolve by plea bargain do not go to trial. The convictions are the result of negotiations in which the central actors are prosecutors, criminal defense attorneys, and defendants. *Id.*

22 *See generally* HANDBOOK OF SOCIAL PSYCHOLOGY (Susan T. Fiske et al. eds., 5th ed. 2010).

23 *See generally* SUSAN T. FISKE & SHELLEY E. TAYLOR, SOCIAL COGNITION (1984).

24 John F. Dovidio, *On the Nature of Contemporary Prejudice: The Third Wave*, 57 J. SOC. ISSUES 829, 829–30 (2001); Phillip A. Goff et al., *The Space Between Us: Stereotype Threat and Distance in Interracial Contexts*, 94 J. PERSONALITY & SOC. PSYCHOL. 91, 92 (2008); Richard T. LaPiere, *Attitudes vs. Actions*, 13 SOCIAL FORCES 230, 232 (1934).

25 *See, e.g.*, THEODOR W. ADORNO ET AL., THE AUTHORITARIAN PERSONALITY (1950); Felicia Pratto et al., *Social Dominance Orientation: A Personality Variable Predicting Social and Political Attitudes*, 67 J. PERSONALITY & SOC. PSYCHOL. 741, 742 (1994).

26 John F. Dovidio et al., *Prejudice, Stereotyping and Discrimination: Theoretical and Empirical Overview*, *in* SAGE HANDBOOK OF PREJUDICE, STEREOTYPING AND DISCRIMINATION 3, 4 (John F. Dovidio et al. eds. 2010); Natasha Warikoo et al., *Examining Racial Bias in Education: A New Approach*, 45 EDUC. RESEARCHER 508, 508–09 (2016).

ences has focused on African American defendants, this Article takes a similar focus out of necessity.

As for the second question, prejudice reduction interventions in domains other than criminal justice have been well studied,[27] and this body of research forms a firm foundation for thinking about the limited interventions that have been tried in criminal justice, as well as those interventions that are currently being recommended. In Part II of this Article, we survey the literature, providing insight into what has been tried, what does and does not appear to work, and what might prove efficacious. Lastly, we develop a novel recommendation that makes use of computational tools as part of an anchoring and vetting process.

I. RACIAL BIAS AND ITS IMPACT ON POST-ARREST AND PRETRIAL DECISION MAKING

In their 1947 book, Allport and Postman reported a widespread cultural stereotype of African Americans as both hot-tempered and likely to use weapons.[28] To illustrate this stereotype, Allport and Postman showed participants a photo in which a white man, situated in a racially diverse group, was holding a knife. After removing the photo from view, Allport and Postman questioned participants about its contents. A significant number stated that the knife was held by a black man, with some even reporting that a black man was brandishing it wildly and threatening white individuals with it.[29] Decades of subsequent research at the nexus of law and psychology has identified stereotypical associations linking blackness with crime, violence, threat, and aggression.[30] In a 2006 experiment that echoed Allport and Postman's work, Payne found stereotypes associating blacks and weapons.[31]

In this Article, our focus is not on the general population. Rather, we focus on prosecutors, defense attorneys, judges, and individuals facing criminal charges. Does racial bias still impact decision making when

---

27 Elizabeth Levy Paluck & Donald P. Green, *Prejudice Reduction: What Works? A Review and Assessment of Research and Practice*, 60 ANN. REV. PSYCHOL. 339, 341 (2009).

28 *See generally* GORDON W. ALLPORT & LEO POSTMAN, THE PSYCHOLOGY OF RUMOR (1947).

29 *Id.* at 111.

30 *See* Joshua Correll et al., *The Police Officer's Dilemma: Using Ethnicity to Disambiguate Potentially Threatening Individuals*, 83 J. PERSONALITY & SOC. PSYCHOL. 1314, 1325 (2003); Jennifer L. Eberhardt et al., *Seeing Black: Race, Crime, and Visual Processing*, 87 J. PERSONALITY & SOC. PSYCHOL. 876, 889 (2004); B. Keith Payne et al., *Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon*, 81 J. PERSONALITY & SOC. PSYCHOL. 181, 189 (2001).

31 B. Keith Payne, *Weapon Bias: Split-Second Decisions and Unintended Stereotyping*, 15 CURRENT DIRECTIONS PSYCHOL. SCI. 287, 287–88 (2006). *See also* B. Keith Payne, Alan J. Lambert & Larry L. Jacoby, *Best Laid Plans: Effects of Goals on Accessibility Bias and Cognitive Control in Race-Based Misperceptions of Weapons*, 38 J. EXPERIMENTAL SOC. PSYCHOL. 384, 384 (2002).

the individual in question is a seasoned district attorney, someone who has a professional obligation for impartial and consistent treatment of individuals of varied race, age, and gender? Does racial bias impact decision making when the individual in question is working on behalf of a minority, as with criminal defense attorneys? Does racial bias infect judges, who are, by nature of their position, supposed to be impartial arbiters? And what is the impact of racial bias on defendants' own decision making, on their willingness to accept a plea bargain? In the succeeding sections, we provide evidence that enables steps in the direction of answering these questions.

A. Prosecutors

Prosecutors are arguably the most influential actors in the U. S. criminal justice system.[32] They have the power to drop a case, to oppose bail or to recommend a certain level of bail, to add or remove charges and counts, to offer and negotiate plea bargains, to propose deals in which a defendant offers testimony against another defendant in exchange for leniency, and to recommend sentences. Prosecutors operate with nearly unchecked discretion and are able to execute the above powers in the absence of meaningful oversight.[33] Thus, when a new case arrives on a prosecutor's desk, he or she has the power to be harsher on certain defendants—and more lenient on others. A 1995 study found that prosecutors are less likely to charge white suspects than minority race suspects.[34] Similarly, Radelet and Pierce reviewed prosecutorial decision making in over 1,000 Florida homicide cases and found a combination of harsher treatment of black defendants and more lenient treatment of

---

[32] *See* Scott A. Gilbert & Molly Treadway Johnson, *The Federal Judicial Center's 1996 Survey of Guideline Experience*, 9 FED. SENT'G REP. 87, 88–89 (1996) (describing district judges' belief that prosecutors' have wide discretion and influence in sentencing); Besiki Kutateladze, Vanessa Lynn & Edward Liang, *Do Race and Ethnicity Matter in Prosecution? A Review of Empirical Studies*, VERA INST. OF JUST. (2012), https://www.vera.org/publications/do-race-and-ethnicity-matter-in-prosecution-a-review-of-empirical-studies. *See generally* Marc L. Miller, *Domination & Dissatisfaction: Prosecutors as Sentencers*, 56 STAN. L. REV. 1211 (2004); Kate Stith, *The Arc of the Pendulum: Judges, Prosecutors, and the Exercise of Discretion*, 117 YALE L.J. 1420 (2008).

[33] *See* Stephanos Bibas, *Prosecutorial Regulation Versus Prosecutorial Accountability*, 157 U. PA. L. REV. 959, 960–61 (2009); Angela J. Davis, *The Prosecution of Black Men*, *in* POLICING THE BLACK MAN: ARREST, PROSECUTION, AND IMPRISONMENT 178, 179 (Angela J. Davis ed., 2017); 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 (2012); Ronald F. Wright, *Elected Prosecutors and Police Accountability*, *in* POLICING THE BLACK MAN: ARREST, PROSECUTION, AND IMPRISONMENT 234, 234–35 (Angela J. Davis ed., 2017). *See generally* ANGELA J. DAVIS, ARBITRARY JUSTICE: THE POWER OF THE AMERICAN PROSECUTOR (2009) [hereinafter DAVIS, ARBITRARY JUSTICE].

[34] Research Working Group of the Task Force on Race and the Criminal Justice System, *Preliminary Report on Race and Washington's Criminal Justice System*, 87 WASH. L. REV. 1, 25 (2012).

white defendants.[35] As Justice Powell acknowledged in *McCleskey v. Kemp*, "[T]he power to be lenient is [also] the power to discriminate."[36]

As tempting as it is to think that each case is unique and that prosecutorial work is a nuanced undertaking, prosecutors often operate with limited information about what actually transpired, and this may lead to reliance on extralegal influences, such as racial stereotypes.[37] An analysis of 2010 and 2011 felony drug cases in New York City found that, of numerous evidentiary factors considered, almost none mattered for predicting sentence offers.[38] In addition to limited information, prosecutors also operate under time constraints, another factor that has been shown to lead to reliance on extralegal factors.[39] Given time and information constraints, legal decision makers may be unable to accurately assess an alleged offender's dangerousness and culpability, and this may make them more likely to rely on stereotypes, such as those that attach to African Americans.[40]

The Department of Justice instructs federal prosecutors, in deciding whether or not to pursue resolution through plea bargaining, to consider a defendant's remorse or contrition and also the interests of the victim, including the victim's right to restitution.[41] Both of these factors are ones likely to be impacted by pervasive racial stereotypes. Perceptions of remorse and contrition are highly subjective, and we know that race may impact related perceptions. For example, across seven studies, Wilson, Hugenberg, and Rule found that people perceive black men as both bigger and more physically threatening than white men, and these misperceptions occurred as the result of both bottom-up cues of racial prototypicality, such as visually perceiving an individual, and top-down information, such as merely being told that an unseen individual belongs

---

35 Michael L. Radelet & Glenn L. Pierce, *Race and Prosecutorial Discretion in Homicide Cases*, 19 LAW & SOC'Y REV. 587, 587 (1985).

36 McCleskey v. Kemp, 481 U.S. 279, 312 (1987) (quoting KENNETH CULP DAVIS, DISCRETIONARY JUSTICE 170 (1969)).

37 *See* Besiki L. Kutateladze, Victoria Z. Lawson & Nancy R. Andiloro, *Does Evidence Really Matter? An Exploratory Analysis of the Role of Evidence in Plea Bargaining in Felony Drug Cases*, 36 LAW & HUM. BEHAV. 431, 439 (2015).

38 *Id.*

39 Celesta A. Albonetti, *An Integration of Theories to Explain Judicial Discretion*, 38 SOC. PROBS. 247, 255 (1991). For a discussion of reliance on heuristics, see DANIEL KAHNEMAN, THINKING, FAST AND SLOW pt. 2 (2011).

40 *See* Darnell F. Hawkins, *Causal Attribution and Punishment for Crime*, 2 DEVIANT BEHAV. 207, 208 (1981); Darrell Steffensmeier, Jeffery Ulmer & John Kramer, *The Interaction of Race, Gender, and Age in Criminal Sentencing: The Punishment Cost of Being Young, Black, and Male*, 36 CRIMINOLOGY 763, 767–68 (1998).

41 U.S. DEP'T OF JUSTICE, JUSTICE MANUAL: PRINCIPLES OF FEDERAL PROSECUTION § 9-27.420 (2018), https://www.justice.gov/jm/jm-9-27000-principles-federal-prosecution#9-27.330.

to a certain racial group.[42] The authors further found that these misperceptions were used to justify beliefs that black men should be subjected to more aggressive law enforcement measures. Goff and colleagues reviewed newspaper reports for 600 cases that took place from 1979-1999 in Philadelphia, finding that, when a defendant was black, there was nearly a fourfold increase in the number of dehumanizing (animalistic) references—and those references strongly correlated with the defendant receiving the most severe punishment available.[43] Whether such beliefs as these affect prosecutors and impact their determinations of remorse or contrition (or even likelihood of recidivism, which is tied up in the same determination), is an empirical question, as is the question of whether defendants are treated more harshly when the defendant/victim dyad is black/white.[44] If prosecutors, as a consequence of such biases, tend to conclude that black males are more intimidating, more aggressive, and perhaps less remorseful for their criminal offenses, and if prosecutors are likely to be harsher on black defendants when the victim is white, then the very guidelines in place for prosecutorial decision making will facilitate racially biased outcomes.

Given this, it is unsurprising that a 2012 meta-analysis of 34 studies found that, in most instances, race affected case outcomes, even after controlling for a host of non-race factors.[45] Prosecutors are less likely to offer blacks a plea bargain, less likely to reduce charge offers for blacks, and more likely to offer blacks plea bargains that include prison time.[46]

---

42 John Paul Wilson, Kurt Hugenberg & Nicholas O. Rule, *Racial Bias in Judgments of Physical Size and Formidability: From Size to Threat*, 113 J. PERSONALITY & SOC. PSYCHOL. 59, 74 (2017).

43 *See* Phillip A. Goff et al., *Not Yet Human: Implicit Knowledge, Historical Dehumanization, and Contemporary Consequences*, 94 J. PERSONALITY & SOC. PSYCHOL. 292, 304 (2008).

44 William J. Bowers, Marla Sandys & Thomas W. Brewer, *Crossing Racial Boundaries: A Closer Look at the Roots of Racial Bias in Capital Sentencing When the Defendant Is Black and the Victim Is White*, 53 DEPAUL L. REV. 1497, 1497 (2004). *See also* Wesley G. Jennings et al., *A Critical Examination of the "White Victim Effect" and Death Penalty Decision-Making from a Propensity Score Matching Approach: The North Carolina Experience*, 42 J. CRIM. JUST. 384, 388 (2014).

45 Kutateladze et al., *supra* note 32, at 7.

46 Celesta A. Albonetti, *Race and the Probability of Pleading Guilty*, 6 J. QUANTITATIVE CRIMINOLOGY 315, 331 (1990); Celesta A. Albonetti, *Charge Reduction: An Analysis of Prosecutorial Discretion in Burglary and Robbery Cases*, 8 J. QUANTITATIVE CRIMINOLOGY 317, 328 (1992); Erika Davis Frenzel & Jeremy D. Ball, *Effects of Individual Characteristics on Plea Negotiations Under Sentencing Guidelines*, 5 J. ETHNICITY IN CRIM. JUST. 59, 59 (2007); Besiki Kutateladze, Nancy R. Andiloro & Brian D. Johnson, *Opening Pandora's Box: How Does Defendant Race Influence Plea Bargaining?*, 33 JUST. Q. 398, 419 (2016) [hereinafter Kutateladze et al, *Opening Pandora's Box*]; Besiki Kutateladze, Nancy R. Andiloro, Brian D. Johnson & Cassia Spohn, *Cumulative Disadvantage: Examining Racial and Ethnic Disparity in Prosecution and Sentencing*, 52 CRIMINOLOGY 514, 514 (2014); John McDonald & Steven Raphael, *An Analysis of Racial and Ethnic Disparities in Case Dispositions and Sentencing Outcomes for Criminal Cases Presented to and Processed by the Office of the San*

Defendants who are black, young, and male fare especially poorly.[47] In a comprehensive study of federal cases, Rehavi and Starr found that blacks received sentences that were 10 percent longer than those of similarly-situated whites and that this disparity was explained by prosecutors' initial charging decisions.[48] Moreover, all else held equal, black arrestees were 75 percent more likely to face a charge with a mandatory minimum sentence than were white arrestees.[49]

Lastly, it is worth noting that explicit racial stereotyping has long been a feature of prosecutorial trial behavior. For years, removal of black jurors was often based on explicit racism. For example, in the 1960s and 1970s, the Dallas County District Attorney's Office used a trial manual that stated, "Do not take Jews, Negroes, Dagos, Mexicans or a member of any minority race on a jury, no matter how rich or how well educated."[50] While the Court ruled in *Batson v. Kentucky*[51] that prosecutors cannot strike a potential juror based solely on race, it is widely acknowledged that prosecutors continue to sidestep the *Batson* Court's admonition.[52] Supreme Court Justice Stephen Breyer wrote that he was not surprised to learn that, despite *Batson*, discriminatory use of peremptory challenges remains common.[53]

B. Criminal Defense Attorneys

Compared with prosecutors, criminal defense attorneys are relatively less well-studied. This scholarly lacuna likely is the result of how defense attorneys are situated in the criminal justice system: they are laboring on behalf of defendants, working to keep them out of prison. Many assume that criminal defense attorneys are unlikely to act in ways that would be detrimental to their clients' interests, which are somewhat

---

Francisco District Attorney, ISSUELAB, Dec. 2017, at 3, http://raceandpolicing.issuelab.org/resources/30712/30712.pdf.

47 Albonetti, *Charge Reduction*, *supra* note 46, at 324; Gail Kellough & Scot Wortley, *Remand for Plea. Bail Decisions and Plea Bargaining as Commensurate Decisions*, 42 BRIT. J. CRIMINOLOGY 186, 196 (2002); Kutateladze et al., *Opening Pandora's Box*, *supra* note 46, at 410.

48 M. Marit Rehavi & Sonja B. Starr, *Racial Disparity in Federal Criminal Sentences*, 122 J. POL. ECON. 1320, 1320 (2014).

49 *Id.* at 1323.

50 Miller-El v. Cockrell, 537 U.S. 322, 333 (2003).

51 Batson v. Kentucky, 476 U.S. 79, 79 (1986).

52 Antony Page, *Batson's Blind-Spot: Unconscious Stereotyping and the Peremptory Challenge*, 85 B.U. L. REV. 155, 235 (2005); Samuel R. Sommers & Michael I. Norton, *Race-Based Judgments, Race-Neutral Justifications: Experimental Examination of Peremptory Use and the Batson Challenge Procedure*, 31 LAW & HUM. BEHAV. 261, 263 (2007).

53 Miller-El v. Dretke, 545 U.S. 231, 266, 268 (2005) (Breyer, J., concurring).

entwinned with the attorneys' own interests.[54] The elected Public Defender of San Francisco said, "My staff works with ethnic minorities every day, I thought. Many had come to this line of work because of concern over racial disparities. How could we be biased against those who we represent?"[55] Given such bias blind spots,[56] the potential effects on case outcomes are large, especially since defense counsel is often the proxy for the defendant for purposes of legal decision-making.[57]

There are reasons to suspect that criminal defense attorneys harbor racial bias. For one, most Americans do,[58] and implicit racial bias affects even those who have positive associations with members of other races.[59] In addition, public defenders are extremely busy, handling on average 371 cases per year, a rate of more than one new case per day.[60] As demands on cognitive resources increase, people are more likely to rely on heuristics, including racial stereotypes.[61]

Criminal defense attorneys are situated similarly to physicians. Physicians are working on behalf of their patients, and they also face time constraints that force them to triage—to carefully allocate their resources

---

54 AMY BACH, ORDINARY INJUSTICE: HOW AMERICA HOLDS COURT 15 (Henry Holt, ed., Holt Paperbacks 2009); Tigran W. Eldred, *Prescriptions for Ethical Blindness: Improving Advocacy for Indigent Defendants in Criminal Cases*, 65 RUTGERS L. REV. 333, 368 (2013).

55 Jeff Adachi, *Public Defenders Can Be Biased, Too, and It Hurts their Non-White Clients*, WASH. POST (June 7, 2016), https://www.washingtonpost.com/posteverything/wp/2016/06/07/public-defenders-can-be-biased-too-and-it-hurts-their-non-white-clients/?utm_term=.fa9d997e2e49.

56 Emily Pronin, Daniel Y. Lin, Lee Ross, *The Bias Blind Spot: Perceptions of Bias in Self Versus Others*, 28 PERSONALITY & SOC. PSYCHOL. BULL. 369, 379 (2002).

57 Molly J. Walker Wilson, *Defense Attorney Bias and the Rush to the Plea*, 65 U. KAN. L. REV. 271, 272 (2016).

58 *See, e.g.,* Kristin A. Lane et al., *Implicit Social Cognition and Law*, 3 ANN. REV. L. & SOC. SCI. 427, 433 (2007); Chris Mooney, *Across America, Whites are Biased and They Don't Even Know It*, WASH. POST (Dec. 8, 2014), https://www.washingtonpost.com/news/wonk/wp/2014/12/08/across-america-whites-are-biased-and-they-dont-even-know-it/?utm_term=.d20002c9de86; Rich Morin, *Exploring Racial Bias Among Biracial and Single-Race Adults: The IAT*, PEW RES. CTR. (Aug. 19, 2015), http://www.pewsocialtrends.org/2015/08/19/exploring-racial-bias-among-biracial-and-single-race-adults-the-iat/.

59 *See, e.g.*, MICHELLE ALEXANDER, THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS 104–09 (2012); Anthony G. Greenwald & Linda Hamilton Krieger, *Implicit Bias: Scientific Foundations*, 94 CAL. L. REV. 945, 952 (2006).

60 SENT'G PROJECT, REPORT OF THE SENTENCING PROJECT TO THE UNITED NATIONS HUMAN RIGHTS COMMITTEE: REGARDING RACIAL DISPARITIES IN THE UNITED STATES CRIMINAL JUSTICE SYSTEM 7 (2013), https://sentencingproject.org/wp-content/uploads/2015/12/Race-and-Justice-Shadow-Report-ICCPR.pdf. *See also* Richard A. Oppel, Jr. & Jugal K. Patel, *One Lawyer, 194 Felony Cases, and No Time*, N.Y. TIMES (Jan. 31, 2019), https://www.nytimes.com/interactive/2019/01/31/us/public-defender-case-loads.html (discussing the heavy caseloads faced by criminal defense attorneys in Louisiana).

61 *See* L. Song Richardson & Phillip Atiba Goff, *Self-Defense and the Suspicion Heuristic*, 98 IOWA L. REV. 293, 304–13 (2012). *See generally* Anuj K. Shah & Daniel M. Oppenheimer, *Heuristics Made Easy: An Effort-Reduction Framework*, 134 PSYCHOL. BULL. 207 (2008).

among a host of needy individuals.[62] Racial bias has been well-documented in physicians.[63] In a study of emergency physicians and patients with serious heart conditions, Green and colleagues found that, as the physicians' pro-white implicit bias increased, so did their likelihood of treating white patients and not treating black patients.[64] Penner and colleagues found that greater implicit bias was negatively associated with physicians' explicit attitudes towards their black patients.[65] Higher prejudice resulted in more problematic interpersonal behaviors on a number of dimensions, including shorter conversation length and diminished patient recall of relayed information.

Three studies have documented racial bias in criminal defense attorneys. Eisenberg and Johnson had 92 capital defense attorneys and 146 "habeas lawyers" (that is, post-conviction lawyers) take a race Implicit Association Test, and the findings showed anti-black implicit bias by white attorneys.[66] A few years later, Edkins also looked at bias in criminal defense.[67] She presented 95 criminal defense attorneys with vignettes in which she manipulated race: "Robert Williams, a 23 year-old African American male" vs. "Robert Williams, a 23 year-old Caucasian American male."[68] She found that plea bargains the defense attorneys recommended that the African American clients accept contained more severe sentences and were more likely to include jail time than did those they recommended for the Caucasian American clients.[69]

The most comprehensive study of criminal defense attitudes was conducted by Avery and colleagues.[70] This survey of 327 criminal defense attorneys practicing in the United States found that, on average, the attorneys surveyed held anti-black implicit bias. When these same attorneys were presented with fictitious clients, the attorneys' implicit bias

---

62 *See* L. Song Richardson & Phillip Atiba Goff, *Implicit Racial Bias in Public Defender Triage*, 122 YALE L.J. 2626, 2632–33 (2013).

63 *See, e.g.*, Janice A. Sabin et al., *Physicians' Implicit and Explicit Attitudes About Race by MD Race, Ethnicity, and Gender*, 20 J. HEALTH CARE FOR POOR & UNDERSERVED 896, 896–97 (2009).

64 Alexander R. Green et al., *Implicit Bias Among Physicians and its Prediction of Thrombolysis Decisions for Black and White Patients*, 22 J. GEN. INTERNAL MED. 1231, 1237 (2007).

65 Louis A. Penner et al., *The Effects of Oncologist Implicit Racial Bias in Racially Discordant Oncology Interactions*, 34 J. CLINICAL ONCOLOGY 2874, 2877 (2016).

66 Theodore Eisenberg & Sheri Lynn Johnson, *Implicit Racial Attitudes of Death Penalty Lawyers*, 53 DEPAUL L. REV. 1539, 1545, 1553 (2004).

67 Vanessa A. Edkins, *Defense Attorney Plea Recommendations and Client Race: Does Zealous Representation Apply Equally to All?*, 35 LAW & HUM. BEHAV. 413, 416 (2011).

68 *Id.* at 417.

69 *Id.* at 419.

70 Joseph J. Avery et al., *Is Your Own Team Against You? Implicit Bias and Interpersonal Regard in Criminal Defense* (2020) (unpublished manuscript) (on file with author).

was correlated with negative explicit attitudes towards these clients—when the clients were black.

Finally, scholars of both law and psychology have argued that criminal defense attorneys are likely to harbor implicit bias,[71] and former criminal defense attorneys, such as Lyon[72] and Rapping,[73] have described racial bias that they observed firsthand. Rankin reported an egregious instance in which a defense attorney who had used racially charged terms in referring to his own client failed to inform the man of a plea bargain that would have allowed him to avoid execution.[74]

C. Judges

Our scope in this Article is limited to post-arrest and pre-trial decision making, and our focus is on plea bargaining, a process in which judges are not typically active.[75] In fact, the role of judges tends to be limited to verifying whether a plea bargain reached was voluntary, knowing, and uncoerced.[76] However, judges still play some role, especially when it comes to sentencing.[77] Prosecutors and defendants and their attorneys bargain outside of court and strike plea deals. Those deals may include negotiated sentences.[78] However, in most instances and in most jurisdictions, the plea bargains struck include only charges and counts, not negotiated sentences.[79] In short, even after the plea, defendants likely face exposure to a range of sentencing outcomes. This is where judges come in. Judges are unlikely to reject the plea bargains that the parties negotiated, but they may impose sentences that are either on the high or low end of what is recommended (or required) by sentencing guidelines.[80]

---

71 *See* Jerry Kang et al., *Implicit Bias in the Courtroom*, 59 UCLA L. REV. 1124, 1141–42 (2012).

72 Andrea D. Lyon, *Race Bias and the Importance of Consciousness for Criminal Defense Attorneys*, 35 SEATTLE U. L. REV. 755, 755–58 (2012).

73 Jonathan A. Rapping, *Implicitly Unjust: How Defenders Can Affect Systemic Racist Assumptions*, 16 N.Y.U. J. LEGIS. & PUB. POL'Y 999, 1000–02 (2013).

74 Bill Rankin, *Racism 'Infected' Killer's Defense? Inmate's Trial to be Scrutinized as Lawyers Ask Board to Commute Osborne's Death Sentence*, ATLANTA J.-CONST., May 30, 2018, at D1. *See also* William S. Sessions, *Death Penalty: Osborne Sentence a Stain on Justice*, ATLANTA J.-CONST., June 3, 2008, at A11.

75 F. Andrew Hessick III & Reshma M. Saujani, *Plea Bargaining and Convicting the Innocent: The Role of the Prosecutor, the Defense Counsel, and the Judge*, 16 BYU J. PUB. L. 189, 221 (2002).

76 *See* Fogus v. United States, 34 F.2d 97, 98 (4th Cir. 1929).

77 Hessick III & Saujani, *supra* note 75, at 237–38.

78 FED. R. CRIM. P. 11 advisory committee's note to 1975 amendment.

79 Hessick III & Saujani, *supra* note 75, at 237–38.

80 Daniel S. McConkie, *Judges as Framers of Plea Bargaining*, 26 STAN. L. & POL'Y REV. 61, 63–64 (2015).

What do we know about racial bias in judges? Rachlinski and colleagues tested 133 U.S. judges on a race Implicit Association Test, and the results showed a strong white preference among white judges and no preference among black judges.[81] In a 2001 study, Mustard examined more than 70,000 federal cases that were sentenced under the Sentencing Reform Act of 1984.[82] After controlling for extensive criminological, demographic, and socioeconomic variables, Mustard found that blacks received substantially longer sentences. In addition, about 56 percent of the black-white disparity was generated by departures from the sentencing guidelines, rather than differential sentencing within the guidelines.[83] Finally, in a dramatic new work, Bielen, Marneffe, and Mocan, working from videos of actual criminal trials, used Virtual Reality technology to replace white defendants in the courtroom with individuals of Middle Eastern or North African descent.[84] Law students, economics students, and practicing lawyers[85] were randomly assigned to watch the trial (with Virtual Reality headsets) from the viewpoint of the judge. These evaluators then made decisions on guilt/innocence and also prison sentence length and fines. The results showed significant overall racial bias against minority defendants in all three: conviction decisions, sentences, and fines.[86]

There is another way beyond sentencing in which judges influence case outcomes: bail decisions. Pretrial detention has been found to exert an influence on defendants' plea bargaining decisions, such that defendants who detained pretrial may plead guilty in order to get out of jail or to decrease their total incarceration time.[87] Heaton, Mayson, and Stevenson analyzed hundreds of thousands of misdemeanor cases from Harris County, Texas, and found that detained defendants were 43% more likely to be sentenced to jail than were similarly-situated non-detained defendants ("releasees").[88] They also received sentences that were more than twice as long as those received by the releasees.[89] Because pretrial deten-

---

81 Jeffrey J. Rachlinski et al., *Does Unconscious Racial Bias Affect Trial Judges?,* 84 NOTRE DAME L. REV. 1196, 1210 (2009).

82 David B. Mustard, *Racial, Ethnic and Gender Disparities in Sentencing: Evidence from the U.S. Federal Courts*, 44 J.L. & ECON. 285, 285–86 (2001).

83 *Id.* at 303.

84 Samantha Bielen et al., *Racial Bias and In-group Bias in Judicial Decisions: Evidence from Virtual Reality Courtrooms*, 43 (Nat'l Bureau of Econ. Res., Working Paper No. 25355, 2018).

85 Because none of the participants were judges, these results can be interpreted as speaking to sentencing decisions generally and not to judicial decision making specifically.

86 Bielen et al., *supra* note 84, at 33.

87 Megan T. Stevenson, *Distortion of Justice: How the Inability to Pay Bail Affects Case Outcomes*, 34 J.L., ECON., & ORG. 511, 512 (2018).

88 Paul Heaton et al., *The Downstream Consequences of Misdemeanor Pretrial Detention*, 69 STAN. L. REV. 711, 717 (2017).

89 *Id.* at 748.

tion typically is the result of defendants' inability to meet bail, racial disparities in detention can be largely explained by the white-black wealth gap. In short, the wealth disparity leads to the pretrial detention disparity which, in turn, distorts case outcomes given the added pressure on detained individuals to plead guilty.[90] Thus, any racial bias that impacts bail decisions will work to compound this already significant disparity.

D. Defendants

The decision making of defendants in accepting or rejecting plea bargains has received far less attention than even criminal defense attorney decision making.[91] This is unsurprising given that social scientific research on bias has tended to focus on the majority group's attitudes and the impact of such attitudes on minorities, with much less emphasis placed on minority groups' own racial attitudes and their impact on behavior and on group dynamics.[92] For our purposes, defendant psychology is of great importance, since defendants have the ultimate power to accept or reject plea offers, and the decision is truly a matter of strategy and of preference.[93]

The dominant theory of plea bargaining is called "bargaining in the shadow of trial," a rational actor model that views a defendant's decision as a tactical one.[94] The calculus is straightforward: if an offered sentence yields value that is greater than the defendant's expected value of going to trial, then the defendant should plead guilty.[95] At its simplest, the defendant would determine the expected sentence if convicted at trial (say, 10 years) and multiply that sentence by the probability of conviction (50%), which would yield a discrete figure (5 years). If the plea bargain is less than that figure, pleading guilty is a rational choice. Research has shown that probabilities of conviction do indeed influence defendants,[96] as do defendants' risk preferences.[97] But the capacity of defendants to

---

90 *See* Stevenson, *supra* note 87, at 512.

91 Rebecca K. Helm & Valerie F. Reyna, *Logical But Incompetent Plea Decisions: A New Approach to Plea Bargaining Grounded in Cognitive Theory*, 23 PSYCHOL., PUB. POL'Y, & L. 367, 376 (2017); *see, e.g.*, Rebecca Hollander-Blumoff, *Social Psychology, Information Processing, and Plea Bargaining*, 91 MARQ. L. REV. 163, 165 (2007).

92 J. Nicole Shelton, *A Reconceptualization of How We Study Issues of Racial Prejudice*, 4 PERSONALITY & SOC. PSYCHOL. REV. 374, 376 (2000).

93 Helm & Reyna, *supra* note 91, at 368.

94 Shawn D. Bushway & Allison D. Redlich, *Is Plea Bargaining in the "Shadow of the Trial" a Mirage?*, 28 J. QUANTITATIVE CRIMINOLOGY 437, 441–42 (2012). *See also* Hollander-Blumoff, *supra* note 91, at 166.

95 Helm & Reyna, *supra* note 91, at 368.

96 *See* Avishalom Tor et al., *Fairness and the Willingness to Accept Plea Bargain Offers*, 7 J. EMPIRICAL LEGAL STUD. 97, 102 (2010).

97 David Bjerk, *On the Role of Plea Bargaining and the Distribution of Sentences in the Absence of Judicial System Frictions*, 28 INT'L REV. L. & ECON. 1, 2 (2008).

plead rationally is disputed, not least because they are often given very little time in which to deliberate.[98] Moreover, factors beyond probability of conviction and risk preference have been shown to be important, including short-term benefits from pleading guilty,[99] although researchers have not unpacked the impact of race on such decisions. For our purposes, the primary point that follows from the "bargaining in the shadow of trial" theory is that, if an African American believes that he or she faces a greater likelihood of conviction at trial, as well as stiffer penalties as the result of this conviction,[100] then he or she will be more likely to accept a suboptimal plea bargain—that is, a plea bargain that would be unacceptable to a similarly-situated white defendant.

E. Interpersonal Interactions and Defendant Decision Making

Racial bias in the United States is significant and persistent, and such bias has a range of consequences, including precipitating negative effects in cognition and physical and mental health.[101] While here we are primarily discussing the effects of whites holding bias, there is an important obverse to the situation: minorities' perceptions of discrimination against them.

Murphy and colleagues examined the cognitive costs of blatant and subtle racial bias during interracial interactions, finding that interacting with a subtly biased white partner impaired blacks' cognitive functioning.[102] Barnes and colleagues found that perceived discrimination is as-

---

98 *See* Tina M. Zottoli et al., *Plea Discounts, Time Pressures, and False Guilty Pleas in Youth and Adults Who Pleaded Guilty to Felonies in New York City*, 22 PSYCHOL. PUB. POL'Y & L. 250, 251 (2016).

99 *See* John H. Blume & Rebecca K. Helm, *The Unexonerated: Factually Innocent Defendants Who Plead Guilty*, 100 CORNELL L. REV. 157, 173 (2014); Tarika Daftary-Kapur & Tina M. Zottoli, *A First Look at the Plea Deal Experiences of Juveniles Tried in Adult Court*, 13 INT'L J. FORENSIC MENTAL HEALTH 323, 331 (2014); Thomas Grisso et al., *Juveniles' Competence to Stand Trial: A Comparison of Adolescents' and Adults' Capacities as Trial Defendants*, 27 LAW & HUM. BEHAV. 333, 357 (2003).

100 *See* John Hagan et al., *Race, Ethnicity, and Youth Perceptions of Criminal Injustice*, 70 AM. SOC. REV. 381, 392–93 (2005); Christopher Muller & Daniel Schrage, *Mass Imprisonment and Trust in the Law*, 651 ANNALS AM. ACAD. POL. & SOC. SCI. 139, 152 (2014).

101 *See* Rodney Clark et al., *Racism as a Stressor for African Americans: A Biopsychosocial Model*, 54 AM. PSYCHOL. 805, 812 (1999); Shelly P. Harrell, *A Multidimensional Conceptualization of Racism-Related Stress: Implications for the Well-Being of People of Color*, 70 AM. J. ORTHOPSYCHIATRY 42, 47–48 (2000); Dorainne J. Levy et al., *Psychological and Biological Responses to Race-Based Social Stress as Pathways to Disparities in Educational Outcomes*, 71 AM. PSYCHOL. 455, 455 (2016); Vetta L. Sanders Thompson, *Racism: Perceptions of Distress Among African Americans*, 38 CMTY. MENTAL HEALTH J. 111, 117 (2002); Shawn O. Utsey & Joseph G. Ponterotto, *Development and Validation of the Index of Race-Related Stress (IRRS)*, 43 J. COUNSELING PSYCHOL. 490, 491 (1996).

102 Mary C. Murphy et al., *Cognitive Costs of Contemporary Prejudice*, 16 GRP. PROCESSES & INTERGROUP REL. 560, 567 (2012).

sociated with worsened cognitive function over time,[103] and Kessler, Mickelson, and Williams found a correlation between perceived discrimination and mental health issues.[104] Sawyer and colleagues found that merely anticipating prejudice leads to stress responses.[105]

But do black defendants perceive that they are discriminated against in U.S. criminal justice? Young and Petersilia wrote that the U.S. criminal justice system has eroded African Americans' rights, creating something of a second-class citizenship.[106] When it comes to encounters with police in particular, white suspects begin such encounters assuming that they have full citizen rights, and they leave the encounters without status loss, but this is not true for black suspects.[107] Indeed, studies have shown that, among African Americans, distrust of criminal justice actors and institutions is significant and longstanding.[108]

Interpersonal interactions between members of different groups are typically highly cognitively demanding, as participants experience anxiety over how they are or may be perceived,[109] with majority group members laboring to appear unbiased[110] and minority group members remaining vigilant for indications that they are being associated with negative stereotypes.[111] Such cognitive demands often result in inter-

---

103 Lisa L. Barnes et al., *Perceived Discrimination and Cognition in Older African Americans*, 18 J. INT'L NEUROPSYCHOLOGICAL SOC'Y 856, 860 (2012).

104 Ronald C. Kessler, Kristin D. Mickelson & David R. Williams, *The Prevalence, Distribution, and Mental Health Correlates of Perceived Discrimination in the United States*, 40 J. HEALTH AND SOC. BEHAV. 208, 223 (1999).

105 *See* Pamela J. Sawyer et al., *Discrimination and the Stress Response: Psychological and Physiological Consequences of Anticipating Prejudice in Interethnic Interactions*, 102 AM. J. PUB. HEALTH 1020, 1024 (2012).

106 *See* Young & Petersilia, *supra* note 11, at 1322.

107 *See* CHARLES R. EPP, STEVEN MAYNARD-MOODY & DONALD P. HAIDER-MARKEL, PULLED OVER: HOW POLICE STOPS DEFINE RACE AND CITIZENSHIP 135–36 (2014).

108 *See* Hagan et al., *supra* note 100, at 381–82; Muller & Schrage, *supra* note 100, at 140. *See generally* DOUGLAS A. BLACKMON, SLAVERY BY ANOTHER NAME: THE RE-ENSLAVEMENT OF BLACK AMERICANS FROM THE CIVIL WAR TO WORLD WAR II (2008); KAHLIL GIBRAN MUHAMMAD, THE CONDEMNATION OF BLACKNESS: RACE, CRIME, AND THE MAKING OF MODERN URBAN AMERICA (2010).

109 *See* J. Nicole Shelton & Jennifer A. Richeson, *Intergroup Contact and Pluralistic Ignorance*, 88 J. PERSONALITY & SOC. PSYCHOL. 91, 91–92 (2005); J. Nicole Shelton, *Interpersonal Concerns in Social Encounters Between Majority and Minority Group Members*, 6 GRP. PROCESSES & INTERGROUP REL. 171, 171 (2003).

110 *See* John F. Dovidio & Samuel L. Gaertner, *Aversive Racism,* 36 ADVANCES EXPERIMENTAL SOC. PSYCHOL. 1, 7 (2004); J. Nicole Shelton et al., *Ironic Effects of Racial Bias During Interracial Interactions*, 16 PSYCHOL. SCI. 397, 397 (2005); Jacquie Vorauer et al., *Meta-Stereotype Activation: Evidence from Indirect Measures for Specific Evaluative Concerns Experienced by Members of Dominant Groups in Intergroup Interaction*, 78 J. PERSONALITY & SOC. PSYCHOL. 690, 690–91 (2000).

111 *See* Nyla R. Branscombe et al., *Perceiving Pervasive Discrimination Among African Americans: Implications for Group Identification and Well-Being*, 77 J. PERSONALITY & SOC. PSYCHOL. 135, 137–38 (1999); Rodolfo Mendoza-Denton et al., *Sensitivity to Status-Based*

group anxiety and attendant behavior,[112] which in turn may be misinterpreted as signs of unfriendliness, dislike, and explicit bias.[113] Moreover, it may lead to communication difficulties and misunderstandings,[114] which is especially concerning in the context of plea bargaining.[115]

In the absence of studies exploring black defendants' plea bargaining decision making, do we have any reason to suspect suboptimal decisions? First, if there are enhanced communication difficulties between minority group clients and majority group criminal defense attorneys, it seems likely that this will result in suboptimal decision making. Second, we know accusatorial interrogation methods increase rates of false confessions.[116] If prosecutors are biased against black defendants and if this bias manifests itself as aggressive and accusatory behavior, then perhaps black defendants will be more likely to accept unfair deals.

The most convincing theory takes up the thread of the above discussion. Many African Americans distrust the criminal justice system,[117] and many perceive bias in criminal justice actors.[118] Their interactions with prosecutors are likely to be stressful and anxiety-inducing, even when the bias prosecutors hold is subtle and inexplicit. Keeping this in mind, we must make clear who pleads guilty and for what reasons. A number of studies have shown that even innocent people are willing to plead guilty.[119] In a 2013 study, Dervan and Edkins found that greater

---

Rejection: Implications for African American Students' College Experience, 83 J. PERSONALITY & SOC. PSYCHOL. 896, 896–97 (2002).

112 *See* Walter S. Stephan & Cookie White Stephan, *An Integrated Threat Theory of Prejudice, in* REDUCING PREJUDICE AND DISCRIMINATION 23, 27 (Stuart Oskamp ed., 2000).

113 *See* Adam R. Pearson et al., *The Fragility of Intergroup Relations: Divergent Effects of Delayed Audiovisual Feedback in Intergroup and Intragroup Interaction*, 19 PSYCHOL. SCI. 1272, 1272 (2008).

114 *See* J. Nicole Shelton & Jennifer A. Richeson, *Interracial Interactions: A Relational Approach,* 38 ADVANCES EXPERIMENTAL SOC. PSYCHOL. 121, 126–27 (2006); J. Nicole Shelton et al., *Prejudice and Intergroup Interaction, in* INTERGROUP MISUNDERSTANDINGS: IMPACT OF DIVERGENT SOCIAL REALITIES 21, 21–22 (Stéphanie Demoulin, Jacques-Philippe Leyens, & John F. Dovidio eds., 2009).

115 *See generally* Stephanos Bibas, *The Myth of the Fully Informed Rational Actor*, 31 ST. LOUIS U. PUB. L. REV. 79, 82 (2011); Allison D. Redlich, *The Validity of Pleading Guilty, in* ADVANCES IN PSYCHOLOGY AND LAW 1, 10–11 (B.H. Bornstein & M.K. Miller eds., 2016); Allison D. Redlich et al., *Understanding Guilty Pleas Through the Lens of Social Science*, 23 PSYCHOL., PUB. POL'Y, & L. 458, 459 (2017).

116 *See* Christian A. Meissner et al., *Accusatorial and Information-Gathering Interrogation Methods and Their Effects on True and False Confessions: A Meta-Analytic Review*, 10 J. EXPERIMENTAL CRIMINOLOGY 459, 460 (2014).

117 See generally BLACKMON, *supra* note 108; MUHAMMAD, *supra* note 108.

118 *See* Hagan et al., *supra* note 100; Muller & Schrage, *supra* note 100.

119 *See, e.g.*, Albert W. Alschuler, *A Nearly Perfect System for Convicting the Innocent*, 79 ALBANY L. REV. 919, 921 (2015); Allison D. Redlich & Reveka V. Shteynberg, *To Plead or Not to Plead: A Comparison of Juvenile and Adult True and False Plea Decisions*, 40 LAW & HUM. BEHAV. 611, 613 (2016); David M. Zimmerman & Samantha Hunter, *Factors Affecting False Guilty Pleas in a Mock Plea Bargaining Scenario*, 23 LEGAL & CRIMINOLOGICAL PSYCHOL. 53, 53 (2018).

than 50% of innocent participants were willing to falsely admit guilt in exchange for a benefit.[120] For the defendant who falsely pleads guilty, the plea may be motivated by a desire to extricate him or herself from the uncomfortable situation.[121] This impulse to overvalue the immediate need—extricating oneself from the uncomfortable situation—is especially pronounced among members of specific populations, including juveniles and individuals with mental illness.[122] A field study of 67 defendants who pleaded guilty found that those defendants who had accepted plea bargains primarily to minimize suffering believed, in retrospect, that the deals they accepted were far from good.[123]

II. INTERVENTIONS

Numerous strategies for reducing bias have been developed, with varying degrees of success.[124] For more than half a century, intergroup contact theory[125] has received wide support in the research community. Contact interventions typically involve participants interacting with outgroup members under optimal conditions (equal group status within the experimental situation, common goals, intergroup cooperation, and authority support),[126] although optimal conditions are not necessarily es-

---

120 Lucian E. Dervan & Vanessa A. Edkins, *The Innocent Defendant's Dilemma: An Innovative Empirical Study of Plea Bargaining's Innocence Problem*, 103 J. CRIM. L. & CRIMINOLOGY 1, 34–35 (2013).

121 Saul M. Kassin et al., *Police-Induced Confessions: Risk Factors and Recommendations*, 34 LAW & HUM. BEHAV. 3, 14 (2010) (discussing false confessions); Lindsay C. Malloy et al., *Interrogations, Confessions, and Guilty Pleas Among Serious Adolescent Offenders*, 38 LAW & HUM. BEHAV. 181 (2014); Allison D. Redlich, *False Confessions, False Guilty Pleas: Similarities and Differences*, *in* POLICE INTERROGATIONS AND FALSE CONFESSIONS: CURRENT RESEARCH, PRACTICE, AND POLICY RECOMMENDATIONS 49 (G. Daniel Lassiter & Christian A. Meissner eds., 2010).

122 Jessica Owen-Kostelnik, Nicholas Dickon Reppucci & Jessica R. Meyer, *Testimony and Interrogation of Minors: Assumptions About Maturity and Morality*, 61 AM. PSYCHOL. 286, 287 (2006); Redlich & Shteynberg, *supra* note 119, at 613; Allison D. Redlich, Alicia Summers & Steven Hoover, *Self-Reported False Confessions and False Guilty Pleas among Offenders with Mental Illness*, 34 LAW & HUM. BEHAV. 79, 81 (2010).

123 Kenneth S. Bordens & John Bassett, *The Plea Bargaining Process From the Defendant's Perspective: A Field Investigation*, 6 BASIC & APPLIED SOC. PSYCHOL. 93, 105 (1985).

124 *See* Calvin K. Lai et al., *Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions*, 143 J. EXPERIMENTAL PSYCHOL.: GEN. 1765, 1779–80 (2014).

125 *See generally* GORDON W. ALLPORT, THE NATURE OF PREJUDICE (1954); ROBIN M. WILLIAMS, JR., THE REDUCTION OF INTERGROUP TENSIONS: A SURVEY OF RESEARCH ON PROBLEMS OF ETHNIC, RACIAL, AND RELIGIOUS GROUP RELATIONS (1947). *See also* John F. Dovidio et al., *Intergroup Contact: The Past, Present, and the Future*, 6 GRP. PROCESSES & INTERGROUP REL. 5, 5 (2003); Thomas F. Pettigrew, *Intergroup Contact Theory*, 49 ANN. REV. PSYCHOL. 65, 75–77 (1998).

126 *See generally* ALLPORT, *supra* note 125; STUART W. COOK ET AL., U.S. DEP'T OF HEALTH, EDUC. AND WELFARE, THE EFFECT OF UNINTENDED INTERRACIAL CONTACT UPON RACIAL INTERACTION AND ATTITUDE CHANGE: FINAL REPORT (1971) (Project No. 5-1320, Cont. No, OEC-4-7-051320-0273), https://files.eric.ed.gov/fulltext/ED060153.pdf.

sential for positive outcomes.[127] Extensions of such interventions may involve having white participants interact with an African American experimenter.[128] Other intergroup interventions include artificially changing group boundaries, such as by sorting participants into non race-based groups.[129]

Many of the different intergroup approaches seem to achieve some measure of short-term success, at least in terms of lower preference for one's in-group.[130] When it comes to feelings and cognitions beyond stated preferences, however, results are decidedly mixed,[131] even though the interventions have been manifold, including explicit instruction for avoiding bias,[132] education concerning social norms,[133] raising awareness of prejudice,[134] forming associations between stigmatized groups

---

127 Thomas F. Pettigrew & Linda R. Tropp, *A Meta-Analytic Test of Intergroup Contact Theory*, 90 J. PERSONALITY & SOC. PSYCHOL. 751, 753 (2006).

128 Brian S. Lowery et al., *Social Influence Effects on Automatic Racial Prejudice*, 81 J. PERSONALITY & SOC. PSYCHOL. 842, 843 (2001).

129 SAMUEL L. GAERTNER & JOHN F. DOVIDIO, REDUCING INTERGROUP BIAS: THE COMMON INGROUP IDENTITY MODEL 63 (Mahzarin Banaji & Miles Hewstone eds., 1st ed. 2000); Richard J. Crisp & Miles Hewstone, *Differential Evaluation of Crossed Category Groups: Patterns, Processes, and Reducing Intergroup Bias*, 2 GRP. PROCESSES & INTERGROUP REL. 307, 325–26 (1999). Richard J. Crisp & Miles Hewstone, *Multiple Social Categorization*, 39 ADVANCES EXPERIMENTAL SOC. PSYCHOL. 163, 236–67 (2007).

130 Paluck & Green, *supra* note 27, at 358.

131 *Id.* at 353.

132 Gareth S. Gardiner, *Complexity Training and Prejudice Reduction*, 2 J. APPLIED SOC. PSYCHOL. 326, 329 (1972); Dan Landis et al., *Can a Black "Culture Assimilator" Increase Racial Understanding?*, 32 J. SOC. ISSUES 169, 180 (1976); Mark Schaller et al., *Training in Statistical Reasoning Inhibits the Formation of Erroneous Group Stereotypes*, 22 PERSONALITY & SOC. PSYCHOL. BULL. 829, 839 (1996).

133 Sheri R. Levy, Steven J. Stroessner & Carol S. Dweck, *Stereotype Formation and Endorsement: The Role of Implicit Theories*, 74 J. PERSONALITY & SOC. PSYCHOL. 1421 (1998); Charles Stangor, Gretchen B. Sechrist & John T. Jost, *Changing Racial Beliefs by Providing Consensus Information*, 27 PERSONALITY & SOC. PSYCHOL. BULL. 486, 487 (2001).

134 Leanne S. Son Hing, Winnie Li & Mark P. Zanna, *Inducing Hypocrisy to Reduce Prejudicial Responses Among Aversive Racists*, 38 J. EXPERIMENTAL SOC. PSYCHOL. 71 (2002).

and positive traits,[135] perspective-taking,[136] and exposure to counter-stereotypical exemplars.[137]

When it comes to long-term success, there is scant evidence that interventions work.[138] Lai and colleagues reviewed nine interventions that had shown short-term efficacy, finding that for none of the nine interventions did the positive effects persist for more than a few days.[139]

There are a number of reasons these interventions show limited efficacy. For one, it is clear that increasing awareness of disparities does not necessarily lead to attitudinal change.[140] Hetey and Eberhardt showed white California voters photographs of incarcerated people in which the racial makeup of the inmates was either 45% or 25% black.[141] When a greater percentage of the inmates were black, the participants were significantly less likely to sign a real petition aimed at lessening racial inequality in prison sentencing.[142] Similarly, when white New York residents read that New York's prison population was 60% black (vs. 40% black), they were less likely to support a petition to end a racially discriminatory NYC policing policy.[143] Peffley and Hurwitz found that, when white participants were informed about racial disparities in execu-

---

135 *See* Bertram Gawronski et al., *When "Just Say No" is Not Enough: Affirmation Versus Negation Training and the Reduction of Automatic Stereotype Activation*, 44 J. EXPERIMENTAL SOC. PSYCHOL. 370, 373 (2008); Michael A. Olson & Russell H. Fazio, *Reducing Automatically Activated Racial Prejudice Through Implicit Evaluative Conditioning*, 32 PERSONALITY & SOC. PSYCHOL. BULL. 421, 422 (2006). *But see* Kerry Kawakami et al., *Just Say No (to Stereotyping): Effects of Training in the Negation of Stereotypic Associations on Stereotype Activation*, 78 J. PERSONALITY & SOC. PSYCHOL. 871, 876 (2000) (demonstrating that negating stereotypes can also reduce stereotype activation).

136 *See* C. DANIEL BATSON, THE ALTRUISM QUESTION: TOWARD A SOCIAL-PSYCHOLOGICAL ANSWER 219–20 (1991). *See generally* Walter G. Stephan & Krystina Finlay, *The Role of Empathy in Improving Intergroup Relations*, 55 J. SOC. ISSUES 729 (1999); Andrew R. Todd et al., *Perspective Taking Combats Automatic Expressions of Racial Bias*, 100 J. PERSONALITY & SOC. PSYCHOL. 1027 (2011). *But see* Jeanine L. Skorinko & Stacey A. Sinclair, *Perspective Taking Can Increase Stereotyping: The Role of Apparent Stereotype Confirmation*, 49 J. EXPERIMENTAL SOC. PSYCHOL. 10, 10 (2013).

137 Nilanjana Dasgupta & Anthony G. Greenwald, *On the Malleability of Automatic Attitudes: Combating Automatic Prejudice with Images of Admired and Disliked Individuals*, 81 J. PERSONALITY & SOC. PSYCHOL. 800, 800–01 (2001).

138 Jeremy Sawyer & Anup Gampa, *Implicit and Explicit Racial Attitudes Changed During Black Lives Matter*, 44 PERSONALITY & SOC. PSYCHOL. BULL. 1039, 1040 (2018); Calvin K. Lai et al., *Reducing Implicit Prejudice*, 7 SOC. & PERSONALITY PSYCHOL. COMPASS 315, 315 (2013).

139 Calvin K. Lai et al., *Reducing Implicit Racial Preferences: II. Intervention Effectiveness Across Time*, 145 J. EXPERIMENTAL PSYCHOL.: GEN. 1001, 1001 (2016).

140 Rebecca C. Hetey & Jennifer L. Eberhardt, *The Numbers Don't Speak for Themselves: Racial Disparities and the Persistence of Inequality in the Criminal Justice System*, 27(3) CURRENT DIRECTIONS PSYCHOL. SCI. 183 (2018).

141 Rebecca C. Hetey & Jennifer L. Eberhardt, *Racial Disparities in Incarceration Increase Acceptance of Punitive Policies*, 25 PSYCHOL. SCI. 1949, 1950 (2014).

142 *Id.* at 1950–51.

143 *Id.* at 1951–52.

tions, 52% favored the death penalty compared with 36% when the participants were not informed of the racial disparities.[144] In short, exposure to disparities may cause individuals to become more supportive of the policies that created those disparities.

Color blindness has been proffered as an appropriate other-regarding stance, although it likewise has proven ineffective at diminishing prejudice.[145] Children who listened to a lesson that contained a color blind message became less likely to alert teachers when genuine racism did occur.[146] Teachers who held color-blind racial attitudes were less willing to adapt their teaching style to minority students' needs,[147] and therapists who held color-blind racial attitudes showed lower empathy towards minority patients.[148] Lastly, whites who read a color-blind passage exhibited more prejudice towards African American and Asian American interaction partners, and this behavior mediated decreased cognitive performance by the minority participants.[149]

A. What's Been Tried: Racial Bias Interventions in Criminal Law

Some of the above-described racial bias interventions have been extended to criminal justice. In order to increase perspective-taking, provide education concerning racial bias, and expose individuals to counter-stereotypical exemplars, Philadelphia police officers have taken trips to the Smithsonian's National Museum of African American History and Culture in Washington, D.C.[150] In 2014, the Manhattan district attorney implemented implicit bias training, which included education, awareness-raising, and explicit instruction.[151] In 2018, the New York Police Department began a training program focused on implicit bias.[152] Critics pointed out that there were no standards for its curriculum and few meth-

---

144 Mark Peffley & Jon Hurwitz, *Persuasion and Resistance: Race and the Death Penalty in America*, 51 AM. J. POL. SCI. 996 (2007).

145 Victoria C. Plaut et al., *Do Color Blindness and Multiculturalism Remedy or Foster Discrimination and Racism?* 27 CURRENT DIRECTIONS PSYCHOL. SCI. 200, 201–02 (2018).

146 Evan P. Apfelbaum et al., *In Blind Pursuit of Racial Equality?* 21(11) PSYCHOL. SCI. 1587, 1590 (2010).

147 Axinja Hachfeld et al., *Should Teachers Be Colorblind? How Multicultural and Egalitarian Beliefs Differentially Relate to Aspects of Teachers' Professional Competence for Teaching in Diverse Classrooms*, 48 TEACHING & TCHR. EDUC. 44, 51 (2015).

148 Alan W. Burkard & Sarah Knox, Effect of Therapist Color-Blindness on Empathy and Attributions in Cross-Cultural Counseling, 51 J. COUNSELING PSYCHOL. 387, 393–94 (2004).

149 Deborah S. Holoien & J. Nicole Shelton, *You Deplete Me: The Cognitive Costs of Colorblindness on Ethnic Minorities*, 48 J. EXPERIMENTAL SOC. PSYCHOL. 562, 564 (2012).

150 Monique Judge, *Philly Police Department Seeks to End Implicit Bias with a Field Trip to a Black Museum*, THE ROOT (July 10, 2018), https://www.theroot.com/philly-police-department-seeks-to-end-implicit-bias-wit-1827493830.

151 *See* James C. McKinley, Jr., *Study Finds Racial Disparity in Criminal Prosecutions*, N.Y. TIMES (July 8, 2014), https://www.nytimes.com/2014/07/09/nyregion/09race.html.

152 Al Baker, *Confronting Implicit Bias in the New York Police Department*, N.Y. TIMES (July 15, 2018), https://www.nytimes.com/2018/07/15/nyregion/bias-training-police.html.

ods for evaluating whether officers successfully implemented the training.[153] Patricia Devine argued that additional data and increasingly precise methods of evaluation are needed to determine if the NYPD intervention has made a difference.[154] This point has been echoed by Paluck and Green: when it comes to interventions for reducing prejudice, the majority of real-world interventions have been studied with nonexperimental methods, and the efficacy of these interventions is difficult to determine.[155]

A novel experimental approach was recently taken by Salmanowitz.[156] Drawing on the academic literature on perspective-taking[157] and virtual reality embodiment,[158] Salmanowitz conducted a virtual reality immersion in which white participants were embodied in black avatars. Compared with a control group that experienced a sham version of the virtual reality program, participants showed significantly lower implicit racial bias.[159] Moreover, in subsequent evaluations of an ambiguous legal case, these same participants produced determinations that were more favorable towards blacks.[160]

Scholars and commentators have recommended racial bias interventions for criminal justice that run the gamut of possibilities. Prominent parties, including *The New York Times* and Seymour W. James Jr., the Attorney-in-Chief at the Legal Aid Society in New York City, have recommended implicit bias trainings for prosecutors.[161] Bennett recommended adding a discussion of implicit bias to jury instructions.[162] Another proposed strategy is to screen potential jurors based on their scores on the Implicit Association Test.[163] Smith and Levinson recom-

---

153 *Id.*

154 *Id.*

155 Paluck & Green, *supra* note 27, at 345.

156 Natalie Salmanowitz, *The Impact of Virtual Reality on Implicit Racial Bias and Mock Legal Decisions*, 5 J.L. & BIOSCIENCES 174, 176 (2018).

157 *See generally* Robyn K. Mallett et al., *Seeing Through Their Eyes: When Majority Group Members Take Collective Action on Behalf of an Outgroup*, 11 GRP. PROCESSES & INTERGROUP REL. 451 (2008).

158 *See generally* Béatrice S. Hasler et al., *Virtual Race Transformation Reverses Racial In-Group Bias*, 12 PLOS ONE e0174965 (2017).

159 Salmanowitz, *supra* note 156, at 194.

160 *Id.* at 199.

161 McKinley, Jr., *supra* note 151; Editorial Bd., *How Race Skews Prosecutions*, N.Y. TIMES (July 14, 2014), https://www.nytimes.com/2014/07/14/opinion/how-race-skews-prosecutions.html.

162 Mark W. Bennett, *Unraveling the Gordian Knot of Implicit Bias in Jury Selection: The Problems of Judge Dominated Voir Dire, the Failed Promise of Batson, and Proposed Solutions*, 4 HARV. L. & POL'Y REV. 149, 169 (2010).

163 *See* Kang et al., *supra* note 71, at 1179–80; Samuel R. Sommers & Phoebe C. Ellsworth, *White Juror Bias: An Investigation of Prejudice Against Black Defendants in the American Courtroom*, 7 PSYCHOL., PUB. POL'Y, & L. 201 (2001).

mended hiring a more diverse pool of assistant district attorneys.[164] The American Bar Association and the National Center for State Courts have been working to educate judges and lawyers about what they can do to mitigate the effects of implicit bias in the courtroom.[165] One possible solution is to place portraits of revered black individuals in U.S. courthouses, which would expose judges, attorneys, and jurors to countertypical exemplars and may create positive associations to replace of existing negative ones.[166]

Perhaps the most consistently cited solution is to blind criminal justice actors to defendants' race.[167] This practice, which is a form of category-masking, is being investigated for potential applications to forensic analysis, expert testimony, and legal fact finding by prosecutors, judges, jurors, and arbitrators.[168] But there are clear limitations: even if a defendant's skin color is hidden from the prosecutor, correlated cues and indicator variables, such as the defendant's zip code, name, or college attended, may still be present.[169]

B. Enlisting Cognitive Computing to Guide Prosecutorial Decision Making

Given the seeming inexorability of racial bias, we are proposing an intervention that focuses more on changing behavior and less on changing attitudes. Specifically, we propose a two-pronged solution. First, prosecutors should be given a race-neutral outcome on which to anchor their decision making. Second, outcomes should be consistently monitored in order to ensure compliance with race-neutral mandates. In this section, we discuss the first prong of this solution, and in the succeeding section we take up the second.

1. Model Creation

The intervention we are proposing focuses on prosecutors. This is not to ignore criminal defense attorneys, judges, and defendants, for whom interventions, if there are any efficacious ones, must come from

---

164 Smith & Levinson, *supra* note 33, at 825.

165 Am. Bar Ass'n & Legal Def. Fund, *Joint Statement on Eliminating Bias in the Criminal Justice System* (July 2015), AM. BAR ASS'N, https://www.americanbar.org/content/dam/aba/images/abanews/aba-ldf_statement.pdf.

166 *See* Dasgupta & Greenwald, *supra* note 137, at 805; Kang et al., *supra* note 71, at 1171.

167 *See* BLINDING AS A SOLUTION TO BIAS: STRENGTHENING BIOMEDICAL SCIENCE, FORENSIC SCIENCE, AND LAW (Christopher T. Robertson & Aaron S. Kesselheim eds., 2016); 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, 69–70 (2015).

168 *See, e.g.*, Sah et al., *supra* note 167.

169 *See* Sunita Sah et al., *Combating Biased Decisionmaking & Promoting Justice & Equal Treatment*, 2 BEHAV. SCI. & POL'Y 78, 84 (2016).

among those described above: implicit bias trainings, perspective taking, increased diversity on public defense staffs—measures that public defenders are working to implement.[170] There are two primary reasons our intervention pertains only to prosecutorial decision making. First, prosecutors hold the majority of power in the PAPT period.[171] Second, it is ironic that the legal actor that has the least direct interaction with the accused—i.e., the prosecutor—is also the most responsible for the accused's fate. This is helpful, however, for forming an efficacious intervention. Because prosecutorial work mainly involves evaluating evidence and making decisions about charges and pleas, an intervention that addresses these objective markers (charges, pleas) and is relatively straightforward to implement, like the one we here propose, stands to yield substantial results.

As discussed above, racial bias—especially implicit racial bias—is persistent in the United States, and most environments exert pressure towards stereotype-consistent behavior. Thus, biased decision making is the default condition in most situations.[172] On account of this default condition operating in the legal sphere, some have recommended using checklists to guide attorney decision making.[173] Others have recommended "nudges" to empower good decisions.[174] Our proposed solution embraces these as foundational while also going further, seeking to use computational analysis to anchor attorney decision making on more just outcomes.

In the U.S. criminal justice system, the general trajectory of a case is thus: a police officer apprehends a suspected malfeasor and files a narrative report of the arrest along with recommended charges. A prosecutor then drafts a formal charging document, which includes criminal charges that may or may not accord with the arresting officer's recommended charges. In some jurisdictions and in some instances, the complaint is reviewed by a grand jury, and the charges are either dismissed or approved through a process called indictment. Finally, the case is resolved, usually by plea bargain. In short, working from the police report, the prosecutor files charges with the aim of reaching a specific outcome, knowing that this outcome will be the result of a compromise with the defendant. For example, after reading the police report and reviewing the

---

170 Adachi, *supra* note 55.

171 *See* Gilbert & Johnson, *supra* note 32, at 89; Miller, *supra* note 32, at 1252; Stith, *supra* note 32, at 1424; Kutateladze et al., *supra* note 32.

172 *See* B. Keith Payne et al., *The Bias of Crowds: How Implicit Bias Bridges Personal and Systemic Prejudice*, 28 PSYCHOL. INQUIRY 233, 242 (2017).

173 *See* Pamela M. Casey et al., *Addressing Implicit Bias in the Courts*, 49 CT. REV. 64, 67 (2013).

174 Ralph Hertwig & Till Grüne-Yanoff, *Nudging and Boosting: Steering or Empowering Good Decisions*, 12 PERSP. PSYCHOL. SCI. 973, 973 (2017).

defendant's basic information and characteristics, a Kings County (New York) district attorney may conclude that the most just outcome would be conviction of a class C felony, grand larceny in the second degree.[175] In order to reach this conviction, however, the defendant will be charged with crimes in addition to grand larceny in the second degree, ones reasonably related to the circumstances of the arrest, and these charges will be dropped during plea bargaining.

In estimating the final disposition of a case, prosecutors have very little on which to base their estimations. There is the arresting officer's narrative and recommended charges; there may be an investigative report; there is basic information about the defendant: race, gender, age, residential address, and criminal record; there is the physical appearance of the defendant in detention, in jail, and/or at arraignment; and there is the prosecutor's past experience, which runs the risk of small sample size bias.

Even though prosecutors are limited in the above-described ways, prosecutors' offices are in possession of information that can be used to form clear and unbiased baselines: hundreds of thousands of closed casefiles. This is the rich set of information from which our solution emanates: using advanced statistical and computer science methods, closed casefiles can be used as the corpus from which to build a model that, based on the arresting officer's narrative report and suggested charges, produces a prediction as to how a case would resolve if the defendant were treated race-neutrally. In essence, the task is a classic machine-learning task: train an algorithm to produce a prediction function that relates case characteristics to case outcomes. This model can then be used to guide prosecutorial decision making: to make it more consistent and less bias prone. The only twist is that the model must be tweaked to account for any observed bias.

At first blush, such a task seems paradoxical: how can a model constructed from biased decision making generate unbiased outcomes? In considering this paradox, it is important to note that, if one were building a model that optimized case outcome predictions, one would want to build a model that accurately captured racial bias. In this instance, bias is not a bug in the system but a feature of the system.[176] After all, machine learning involves using feature frequency information to statistically map correspondence between features and outcomes. Indeed, this approach, in

---

175 N.Y. Penal Law §155.40(1) (McKinney 2019).

176 Danny Li, *AOC Is Right: Algorithms Will Always Be Biased As Long As There's Systemic Racism in This Country*, SLATE (Feb. 1, 2019), https://slate.com/news-and-politics/2019/02/aoc-algorithms-racist-bias.html. (arguing that, "While algorithmic processes are intended to rationalize decision-making by minimizing human bias and fallibility, the data they consume already reflects that bias, and so algorithms continue to perpetuate racial bias, just at a greater speed and with the illusion of independence.").

which bias is a feature, is the usual one when it comes to predictive analytics in the law. For example, statistical algorithms are currently used for risk assessment, and some of these resultant tools have been charged with perpetuating racial bias[177] and have been the subject of litigation.[178]

Our proposed intervention is distinct from these approaches in that the goal is to remove suspect features, such as racial bias, from the model, thus bringing its predictions into closer accord with Constitutional mandates for racially equal treatment of criminal defendants by state actors. Its job is to suggest Constitutionally proper outcomes. For this reason, we are calling the model not a predictive model but a "suggestive model."

The suggestive modeling approach has scholarly precedents. Broadly, we can look to Prasad and colleagues, who, drawing on suboptimal historical patient treatment profiles, have used reinforcement learning algorithms to determine optimal treatment for patients in intensive care units.[179] Likewise, Mancuhan and Clifton[180] focused on discovering discrimination in data sets and preventing discrimination in subsequent classification. They measured the effect of a protected attribute, such as race, in a subset of the dataset using an estimated probability distribution via a Bayesian network. Then they proposed a classification method that corrected for the discovered discrimination without using the protected attribute (race) in the decision process.[181]

Along these lines, Lipton and colleagues suggested that a weak version of disparate treatment can involve accessing the protected attribute (race) during training but omitting the attribute during classification.[182] That is, the attribute is accessed during the inductive step and elided during the deductive step. Zliobaite and Custer also employ race in training a model that then avoids discrimination at classification.[183] These approaches build from the recognition, long established in the scholarly

---

177 Julia Angwin et al., *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. *See also* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 674 (2016).

178 *See* State v. Loomis, 881 N.W.2d 749, 753 (Wis. 2016).

179 Niranjani Prasad et al., *A Reinforcement Learning Approach to Weaning of Mechanical Ventilation in Intensive Care Units*, UNCERTAINTY ARTIFICIAL INTELLIGENCE (2017), http://auai.org/uai2017/proceedings/papers/209.pdf.

180 *See* Koray Mancuhan & Chris Clifton, *Combating Discrimination using Bayesian Networks*, 22 ARTIFICIAL INTELLIGENCE L. 212, 212 (2014).

181 *See id.* at 219.

182 Zachary C. Lipton et al., *Does Mitigating ML's Disparate Impact Disparity Require Treatment Disparity?*, ARXIV 2 (2017), https://arxiv.org/abs/1711.07076.

183 Indrė Žliobaitė & Bart Custers, *Using Sensitive Personal Data May be Necessary for Avoiding Discrimination in Data-driven Decision Models*, 24 ARTIFICIAL INTELLIGENCE L. 183, 183 (2016).

community, that not only does blindness not entail fairness,[184] but that it also constitutes a poor notion of fairness.[185]

If the data are proportionately biased, such that there is, for example, an underrepresentation of white drug offenders, techniques such as one developed by Amini and colleagues are useful.[186] These researchers created an algorithm that fuses an original learning task with a variational autoencoder in order to learn the latent structure within the dataset. The algorithm then adaptively uses the learned latent distributions to reweight the importance of certain features during training. The researchers have used this algorithm to address racial and gender bias in facial detection systems.[187]

Computational techniques available for suggestive model creation are manifold,[188] and full discussion of them is beyond the scope of this Article,[189] but the model should be tested via tasks that demand classification of individual cases in a sample independent of the data set on which the model was originally trained.[190] The model also should be used with active cases—simultaneous to prosecutorial handling of these cases—to evaluate machine performance against prosecutor performance, focusing on predictive accuracy and also relative racial bias. If necessary, the model should be reweighted according to testing results. Indeed, testing, modifying, and validating are central tasks in the field of

---

184 Cynthia Dwork et al., *Fairness Through Awareness,* PROC. 3RD INNOVATIONS THEORETICAL COMPUT. SCI. CONF. 214, 218 (2012), https://arxiv.org/pdf/1104.3913.pdf.

185 Moritz Hardt, Eric Price & Nathan Srebro, *Equality of Opportunity in Supervised Learning,* October 11, 2016, ARXIV, https://arxiv.org/pdf/1610.02413.pdf.

186 Alexander Amini et al., *Uncovering and Mitigating Algorithmic Bias through Learned Latent Structure*, Ass'n Advancement Artificial Intelligence/Ass'n Computing Mach. Conf. Artificial Intelligence, Ethics, and Soc'y 289, 289 (2019), http://www.aies-conference.com/wp-content/papers/main/AIES-19_paper_220.pdf. For another approach at building a non-discriminatory classifier, see Irene Chen et al., *Why Is My Classifier Discriminatory?*, ARXIV 1 (2018), https://arxiv.org/pdf/1805.12002.pdf.

187 *See* Amini et al., *supra* note 186.

188 *See, e.g.*, *AI Fairness 360 Open Source Toolkit*, http://aif360.mybluemix.net/?utm_campaign=the_algorithm.unpaid.engagement&utm_source=hs_email&utm_medium=email&utm_content=69523284&_hsenc=P2ANqtz-8u7vGSnvu62JcEEIzyy1ZRkmBf7RHOyJoQJyihXQYAQRUVNStkVlgOxHxj9alH-ojuwGu6iSJ1LwP8TNVKBIXN24SLVw&_hsmi=69523284 (IBM has compiled an open source toolkit that contains over 70 fairness metrics and 10 state-of-the-art bias mitigation algorithms developed by the research community).

189 *See* Alvin Rajkomar et al., *Scalable and Accurate Deep Learning with Electronic Health Records*, NPJ DIG. MED. 1 (2018) (providing a successful predictive approach in a slightly different domain).

190 *See, e.g.*, Daniel Martin Katz et al., *A General Approach for Predicting the Behavior of the Supreme Court of the United States*, 12 PLOS ONE e0174698 (2017) (discussing out-of-sample testing); Tal Yarkoni & Jacob Westfall, *Choosing Prediction Over Explanation in Psychology: Lessons from Machine Learning*, 12 PERSPECTIVES PSYCHOL. SCI. 1100, 1112 (2017) (discussing k-fold cross validation).

machine learning[191] and should be essential components of model construction.

The completed model will enable introduction, early in the development of a new criminal case, of an anchor to prosecutorial decision making. That is, the model's race-neutral recommendation will be used by the prosecutor as an anchor; if he or she deviates from the model's prediction, the reason for the deviation must be documented by the prosecutor and used to assess model performance moving forward.

2. The Human Element

This is not a model to replace prosecutors. Rather, it should form half of a computer-human problem-solving pair.[192] While machines can be more consistent than human actors, and while they possess superior memory and computational capabilities, they lack judgment and may struggle with outliers.[193] For this reason, the model should complement prosecutors and form part of a collaborative problem-solving pair, augmenting human capabilities rather than replacing or obviating them.[194]

Within this problem-solving pair, the machine plays the role of prediction: it provides the baselines. The prosecutor plays the judgment role: he or she makes legal decisions given the full range of inputs, which includes the baseline provided by the machine. This teamwork approach is currently in operation at Vanguard, the investment management company.[195] There, cognitive technology is used to generate the financial plan and to provide goal-based forecasting in real time. The human adviser is tasked with understanding the investment goal, customizing the plan based on the particularities of the client, and serving as an interactive coaching partner. Similarly, in our cognitive technology/prosecutor pairing, the prosecutor is tasked with understanding prosecution goals, customizing the plan based on case particularities, and interacting with defendants, the defense attorney, and the judge.

3. Ensemble Modeling

Rather than rely on a single computational model, prosecutors' offices should introduce and compare models with different theoretical un-

191 *See* Jon Kleinberg et al., *Human Decisions and Machine Predictions*, 133 Q. J. ECON. 237 (2018).

192 *See generally* KEVIN D. ASHLEY, ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS: NEW TOOLS FOR LAW PRACTICE IN THE DIGITAL AGE (2017); JOHN KELLY III & STEVE HAMM, SMART MACHINES: IBM'S WATSON AND THE ERA OF COGNITIVE COMPUTING (2013); Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87 (2014).

193 *See* KELLY III & HAMM, *supra* note 192, at 27–28.

194 Thomas H. Davenport & Rajeev Ronanki, *Artificial Intelligence for the Real World*, HARV. BUS. REV., Jan.–Feb. 2018, at 108, 111.

195 *Id.* at 115.

derpinnings and aims. This process of triangulating among different models can be called "ensemble modeling," and it is the appropriate approach in a domain like criminal justice, where the goal is more nuanced than mere prediction maximization.

Each prosecutor's office needs to clarify its own priorities. Cognitive computing should not be used to solve human problems until humans have figured out how to solve those problems themselves. Cognitive computing is helpful for scaling solutions and for regularizing solutions once found—and this is how we are proposing that it be used. But the priorities must be ironed out by district attorneys and other decision makers. For this, we recommend that each office create a reasoned rule.[196]

Developing a reasoned rule would involve selecting a few variables that are incontrovertibly related to an outcome (such as final plea bargain charges, or whether a defendant should receive a charge reduction from arrest to final disposition). These variables should be assigned equal weight in the prediction formula, and their valence should be positive or negative depending on the specific variable. For example, a history of violent felony convictions might be negative (less likely to result in a charge reduction), while a primary charge of a drug crime might be positive (more likely to result in a charge reduction). Jongbin Jung and colleagues emphasized that such rules are valuable because they, one, enable decisions that can be made quickly and without a computer; two, require only limited information to reach decisions; and, three, allow for insight into justifications for the decisions made.[197] Reasoned rule models have proven successful in judicial decision making,[198] personnel selection, election forecasting, and many other tasks.[199]

One problem that tracks the use of cognitive computing in public domains is interpretability. There is general resistance to machine learning outcomes when those outcomes are determined by inscrutable processes.[200] Thus, it is important that suggestions produced by our full model (described in the previous section) be discussed in relation to suggestions produced by a reasoned rule model. In addition, computational techniques should be used to increase the transparency of the full model itself. We recommend local-interpretable-model agnostic explanations,

---

196 *See* Daniel Kahneman, Andrew M. Rosenfield, Linnea Gandhi & Tom Blaser, *Noise: How to Overcome the High, Hidden Cost of Inconsistent Decision Making*, HARV. BUS. REV., Oct. 2016, at 38, 44–45.

197 *See* Jongbin Jung et al., *Creating Simple Rules for Complex Decisions*, HARV. BUS. REV., Apr. 2017, at 1.

198 *See id.*

199 *See* Kahneman et al., *supra* note 196.

200 *See* Li, *supra* note 176.

which identify the aspects of input data on which a trained model relies as it makes its predictions.[201]

4. Metrics of Success

The efficacy of the proposed model is an important consideration. The model will produce a suggested case outcome, and the efficacy of the intervention depends on the extent to which this nonbinding suggestion influences prosecutorial decision making. In short, will such an anchor be influential?

In their classic 1974 *Science* article, Tversky and Kahneman identified the human tendency to make estimates by starting from an initial value—which may or may not be rationally related to a given problem—and adjusting towards a final answer.[202] These adjustments are often insufficient.[203] Tversky and Kahneman called this phenomenon "anchoring."[204] A boat chained to an anchor can drift only so far. The anchoring effect has been extensively researched in fields as diverse as economics,[205] medicine,[206] and law,[207] and it has been observed in contexts that are not strictly numerical, as with estimations of physical size.[208] Although much scholarly effort has been invested in mitigating anchoring biases,[209] some scholars have considered how anchoring might be used to precipitate desirable behaviors.[210]

An anchor, on its own, might not be enough to change behavior. In Kentucky, pretrial risk assessments had only a small and temporary effect on judges' decisions, and the assessments did not lessen racial dis-

---

201 *See* Michael Chui et al., *What AI Can and Can't Do Yet for Your Business*, MCKINSEY Q. 96 (Jan. 2018), https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/what-ai-can-and-cant-do-yet-for-your-business.

202 *See* Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCI. 1124, 1128 (1974).

203 *See* Paul Slovic & Sarah Lichtenstein, *Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment*, 6 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 649, 701 (1970).

204 Tversky & Kahneman, *supra* note 202, at 1128.

205 Sean D. Campbell & Steven A. Sharpe, *Anchoring Bias in Consensus Forecasts and its Effect on Market Prices*, 44 J. FIN. & QUANT. ANALYSIS 369, 388 (2009).

206 Andi L. Foley & Jennifer James-Salwey, *Stigma, Anchoring, and Triage Decisions*, 42 J. EMERGENCY NURSING 87, 87 (2016).

207 Mark W. Bennett, *Confronting Cognitive 'Anchoring Effect' and 'Blind Spot' Biases in Federal Sentencing: A Modest Solution for Reforming a Fundamental Flaw*, 104 J. CRIM. L. & CRIMINOLOGY 489, 492 (2014).

208 Robyn A. LeBoeuf & Eldar Shafir, *The Long and Short of It: Physical Anchoring Effects*, 19 J. BEHAV. DECISION MAKING 393, 404 (2006).

209 *See* Bradley J. Adame, *Training in the Mitigation of Anchoring Bias: A Test of the Consider-the-Opposite Strategy*, 53 LEARNING & MOTIVATION 36, 38 (2016).

210 *See* Elias Oussedik et al., *An Anchoring-Based Intervention to Increase Patient Willingness to Use Injectable Medication in Psoriasis*, 153 JAMA DERMATOLOGY 932, 933 (2017).

parity in detentions.[211] In Chicago, algorithms designed to improve equity in bond decisions had some effect, but they did not reliably change judges' behavior.[212]

Thus, we propose two additional checks. First, if a prosecutor deviates significantly from the suggested outcome, he or she should document the reasons for the deviation. Past research has emphasized that human justification of deviations significantly impacts behavior within human-machine interactions, not least because deviations create more work for the decision maker.[213] In addition to creating documentation of the prosecutor's reasoning, this also will help slow the prosecutor's decision making, which may on its own help to alleviate the tendency to rely on heuristics, such as racial stereotypes.[214] The second check of the efficacy of our anchoring model involves more robust data collection, analysis, and communication of results to prosecutors, which is discussed in Section II.C.

It is not enough to determine whether the model is improving case fairness. We also must ask two additional questions. First, is the model decreasing case outcome noise?[215] While we have been discussing bias in this Article, prosecutors' decisions can be faulty for a different reason: they may have high variance, what is called "noise." Noise occurs for a number of reasons, such as a prosecutor's last case seen. A prosecutor who has just processed a murder case might treat a robbery differently than if the same prosecutor had just processed a drug case. In this way, there is variance across occasions. There also will be variance across prosecutors. Some might be systematically harsher than others.[216] By implementing a model like the one we propose, not only will bias be addressed, noise also will be decreased. The baselines produced will help to ensure consistency across both occasions and persons.

---

211 *See* Megan Stevenson, *Assessing Risk Assessment in Action*, 103 MINN. L. REV. 303, 370 (2018).

212 *See* The Coalition to End Money Bond, *Monitoring Cook County's Central Bond Court: A Community Courtwatching Initiative* (Feb. 27, 2018), http://www.chicagoappleseed.org/wp-content/uploads/2018/02/Courtwatching-Report_Coalition-to-End-Money-Bond_FINAL_2-25-18.pdf.

213 *See* Arno Lodder & Anja Oskamp, *Law, Information Technology, and Artificial Intelligence, in* INFORMATION TECHNOLOGY AND LAWYERS: ADVANCED TECHNOLOGY IN THE LEGAL DOMAIN, FROM CHALLENGES TO DAILY ROUTINE 1 (Anja Oskamp & Arno Lodder eds., 2006).

214 *See generally* KAHNEMAN, *supra* note 39.

215 Kahneman et al., *supra* note 196, at 40–42.

216 *See, e.g.*, BRUCE FREDERICK & DON STEMEN, THE ANATOMY OF DISCRETION: AN ANALYSIS OF PROSECUTORIAL DECISION MAKING – TECHNICAL REPORT, at 1, 48 (2012), https://www.ncjrs.gov/pdffiles1/nij/grants/240334.pdf (noting that, in a large county prosecutor's office studied by the authors, a major difficulty in hiring new prosecutors was figuring out how to identify in advance those "prosecutors who would tend to be overly punitive or harsh," "go[ing] to the max on every case").

The second question is whether the model is increasing prosecutorial efficiency. On the one hand, when a prosecutor disagrees with the model's suggested outcome, there will be a cost in time, as the prosecutor will need to document the reason for the deviation. On the other hand, because a baseline outcome will be suggested for each case, prosecutors should be able to process routine cases faster. If a case is unusual—say, a rare crime, or even a typical crime with atypical evidence—then prosecutors will have to be spend more time on the case, but the time spent should be no greater than what they would have spent in the absence of the model. The time savings across cases should be tallied, and those savings should be reinvested. They could be used to increase time spent on the outlier cases (the unusual cases). The savings also could be used to calculate lower staffing needs, such that an office might not need to hire additional prosecutors in future years, thus freeing up funds to be reinvested in diversionary programs or social services or anything else that might benefit defendants.

5. Similar Attempts and Additional Implications

One prominent human-machine collaboration is ongoing in social services, where machine learning is being employed to cluster incoming reports of child abuse into low- or high-risk categories.[217] The social workers who receive these calls/reports are given final say (as are prosecutors in our proposed model), but they are encouraged to use the computer predictions to inform their thinking. Recent evaluation reveals that the computers are outperforming the human evaluators, making more true negative identifications and fewer false positive ones.[218]

Other prominent collaborations are ongoing in the judiciary, where the potential for impactful human-machine collaborations is significant.[219] The Public Safety Assessment tool (PSA) is designed to reduce bias in judicial decision making. In assessing dangerousness, the tool takes into account age and history of criminal convictions, but it elides race, gender, employment background, where a defendant/offender lives, and history of criminal arrests.[220] Similarly, risk assessments focusing on

---

217 *See* Dan Hurley, *Can an Algorithm Tell When Kids Are in Danger?*, N.Y. TIMES MAG. (Jan. 2, 2018), https://www.nytimes.com/2018/01/02/magazine/can-an-algorithm-tell-when-kids-are-in-danger.html.

218 *See id.*

219 *See, e.g.*, Kleinberg et al., *supra* note 191, at 238; GRANT T. HARRIS ET AL., VIOLENT OFFENDERS: APPRAISING AND MANAGING RISK, 5–7 (3d ed. 2015); Richard P. Kern & Meredith Farrar-Owens, *Sentencing Guidelines with Integrated Offender Risk Assessment*, 25 FED. SENT'G REP. 176, 177 (2013).

220 Arnold Foundation, *Public Safety Assessment: Risk Factors and Formula* (2018), https://www.psapretrial.org/about/factors.

recidivism are frequently consulted by sentencing courts.[221] The Wisconsin Supreme Court[222] has held that judges may use such computer aided risk scores, although it cannot be their sole consideration.[223]

We must be mindful that technical answers cannot resolve moral questions. There are a host of issues attendant to our proposed intervention that must be discussed with policy makers, prosecutors, and key criminal justice actors. There are live moral issues that need to be addressed openly and with philosophical seriousness. One approach to such issues is deliberative polling, which brings together a representative group of residents in a community to debate policy goals and reach consensus.[224]

Consider a few of the issues. First, there is tension in choosing between group and individual fairness. The group fairness impossibility theorem stems from the practical matter that different groups almost always have unequal base rates.[225] Given this, if one has a calibrated score (for whatever potential event), and one picks the same threshold for all groups, the error rates will be unequal. In other words, one must choose between individual-level fairness and group-level fairness.

Second, given that criminal histories are an accepted component in sentencing decisions, it makes sense that our proposed model would draw on them in making charge and sentence recommendations. Will this lead to an unfair increase in racial disparities in incarceration? Consider the parallel case of mandatory minimum sentences, which often lead to an increase in racially disparate legal outcomes, and where this effect tends to be driven less by racial bias and more by the fact that minorities, on average, have more substantial criminal histories.[226] If we think that criminal history—especially history of arrests—is the product of historical bias and oppression, then a conundrum emerges: should the model

---

221 *See* Timothy Bakken, *The Continued Failure of Modern Law to Create Fairness and Efficiency: The Presentence Investigation Report and Its Effect on Justice*, 40 N.Y.L. SCH. L. REV. 363, 370 (1996). *See also* Sonja B. Starr, *Evidence Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803, 816 (2014) (discussing risk assessments at sentencing).

222 State v. Loomis, 881 N.W.2d 749, 757 (Wis. 2016).

223 Note, State v. Loomis*: Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessments in Sentencing*, 130 HARV. L. REV. 1530, 1532 (2017).

224 *See* Christian List, Robert C. Luskin, James S. Fishkin & Iain McLean, *Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls*, 75 J. POL. 80, 85 (2012).

225 *See* Alexandra Chouldechova, *Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments*, 5 BIG DATA 153, 153 (2017). *See also* Jon Kleinberg et al., *Inherent Trade-Offs in the Fair Determination of Risk Scores*, PROC. INNOVATIONS THEORETICAL COMPUT. SCI. (2017), https://arxiv.org/abs/1609.05807.

226 *See* Joshua B. Fischman & Max M. Schanzenbach, *Racial Disparities Under the Federal Sentencing Guidelines: The Role of Judicial Discretion and Mandatory Minimums*, 9 J. EMPIRICAL LEGAL STUD. 729, 737 (2012).

disregard criminal history for those who have previously engaged with the legal system and use it only in instances of first offense, that is, use it only from this point forward? This, of course, would result in dissimilarly situated defendants (one with, and one without, a record) receiving the same treatment, a possibility that itself suggests unequal treatment.

Third, there is the question of transparency. Much of the litigation surrounding machine assessment tools concerns the proprietary nature of the algorithms that produce the assessments.[227] For our proposed intervention, transparency is compulsory. The algorithms developed must be publicly disclosed, as should the machine learning methods used and the assumptions that inform the use of these methods. As is being shown in recent initiatives, including one in New York City,[228] openness can increase the legitimacy of and public confidence in the legal system. Or consider the Allegheny Family Screening Tool. This predictive-risk modeling tool was created to help improve child welfare screening decisions.[229] It is owned by the county. Its workings are public. Its criteria are described in academic publications and discussed by local officials. At public meetings held in Pittsburgh before the system's adoption, lawyers, child advocates, parents, and even former foster children asked hard questions of both the academics and the county administrators.[230]

C. Standardizing and Improving Data Collection and Analysis

Building a computational model that performs well and that minimizes expected prediction error typically requires large amounts of data.[231] Prosecutors' offices have datasets that include thousands and even hundreds of thousands of cases. In conversations with prosecutors, we have learned that this data is in varying states of order and disorder. Some offices have internal divisions that have organized decades of casefiles, rendered them into manageable electronic formats, and actively

---

227 *See* State v. Loomis, 881 N.W.2d 749, 765 (Wis. 2016).

228 *See* Julia Powles, *New York City's Bold, Flawed Attempt to Make Algorithms Accountable*, NEW YORKER (Dec. 20, 2017), https://www.newyorker.com/tech/annals-of-technology/new-york-citys-bold-flawed-attempt-to-make-algorithms-accountable.

229 *See* Allegheny Cty., *Predictive Risk Modeling in Child Welfare in Allegheny County: The Allegheny Family Screening Tool* (2018), https://www.alleghenycounty.us/Human-Services/News-Events/Accomplishments/Allegheny-Family-Screening-Tool.aspx; 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.

230 Allegheny Cty., *supra* note 229.

231 *See* TREVOR HASTIE, ROBERT TIBSHIRANI & JEROME FRIEDMAN, THE ELEMENTS OF STATISTICAL LEARNING (2009); Pedro Domingos, *A Few Useful Things to Know about Machine Learning*, 55 COMM. ACM 78, 84 (2012); Daniel M. McNeish, *Using Lasso for Predictor Selection and to Assuage Overfitting: A Method Long Overlooked in Behavioral Sciences*, 50 MULTIVARIATE BEHAV. RES. 471 (2015); Tal Yarkoni & Jacob Westfall, *Choosing Prediction over Explanation in Psychology: Lessons from Machine Learning*, 12 PERSPECTIVES PSYCHOL. SCI. 1100 (2017).

update the files. Some are on the other end of the spectrum, with only cases from recent years stored electronically, and even these are stored haphazardly. Building the proposed model will require more offices to update their data capabilities, and prosecutors' offices must function more like large private law firms and corporations, which keep detailed and extensive records.

One strand of scholarship proposes institutional controls to correct for improper use of discretion by prosecutors. Vorenberg, for example, argued for legislative restructuring of the prosecutorial system.[232] Others have called for review of individual cases.[233] Such reforms, however, to the extent that they have been attempted, have not worked, as legislative restructuring is too expansive an undertaking and individualized review too narrow.[234] We must remember that the last time a claim of racially selective prosecution was successful was 1886.[235]

The justice system at present is one that prioritizes prosecutorial discretion but is characterized by opacity and insularity.[236] Our suggestion, which works in tandem with the full computational model described in the preceding sections of this Article, demands the opposite: there should be regular and open statistical review of prosecutorial workings. Such reviews should be quarterly, or biannual, and they should make clear to both prosecutors and the public the extent to which similarly-situated defendants are being treated similarly.

These reviews will serve three primary purposes. First, they will act to vet the efficacy of the model proposed in the succeeding section. Is the model working? Is it changing attorney decision making? Are similarly-situated defendants now more likely to be treated similarly? Second, we suspect that even well-meaning prosecutors, acting on implicit cognitions, may treat defendants unfairly, and they may not be aware that they are doing so. Regular reviews will provide internal feedback to improve behavior, both at the individual prosecutor level and at the broader of-

---

232 *See* James Vorenberg, *Decent Restraint of Prosecutorial Power*, 94 HARV. L. REV. 1521, 1566 (1981). *See also* Arnold I. Burns, Warren L. Dennis & Amybeth Garcia-Bokor, *Curbing Prosecutorial Excess: A Job for the Courts & Congress*, CHAMPION (1998), https://www.nacdl.org/champion/articles/98jul01.htm; Daniel C. Richman & William J. Stuntz, *Al Capone's Revenge: An Essay on the Political Economy of Pretextual Prosecution*, 105 COLUM. L. REV. 583, 583 (2005).

233 DAVIS, ARBITRARY JUSTICE, *supra* note 33, at 3–18; Richard A. Rosen, *Disciplinary Sanctions Against Prosecutors for Brady Violations: A Paper Tiger*, 65 N.C. L. REV. 693, 697 (1987).

234 Bibas, *supra* note 33, at 962.

235 *See generally* Yick Wo v. Hopkins, 118 U.S. 356 (1886) (ruling that selective enforcement of fire safety ordinances against Chinese Americans was unconstitutional).

236 Stephanos Bibas, *Plea Bargaining outside the Shadow of Trial*, 117 HARV. L. REV. 2464, 2500 (2004).

fice, city, county, and state levels.[237] Some have argued that prosecutors who are guided by data will be less likely to rely on racial biases.[238]

Third, and perhaps most importantly, regular and open reviews will increase trust in the criminal justice system. As discussed above, defendants' attitudes towards prosecutors, criminal defense attorneys, and the criminal justice system itself are important. A system that yields fair results and also clear reports outlining those results is the surest way to increase belief in, reliance on, and adherence to that system.

Currently, there are resources that provide access to criminal justice data. Measures for Justice, for example, was founded in 2011 to develop data-driven measures to assess the criminal justice process on a county-by-county basis.[239] There also is the Uniform Crime Reporting Program Data Series, which is compiled by the Federal Bureau of Investigation,[240] and there is data made available by the United States Sentencing Commission, which focuses on federal cases.[241] Some governmental agencies, such as the Bureau of Justice Statistics, produce reports analyzing some of this data.

However, what is needed is data from the individual offices. Such data exists—the only question is what state the data is in. In line with our recommendation, a few prosecutors' offices around the U.S. have begun organizing their institutional data.[242] Similar efforts at standardizing legal data collection are underway across the European Union[243] and Brazil.[244] Projects like these are precisely the way forward.

CONCLUSION

Racial bias that emerges post-arrest and pretrial is vitally important given that almost all criminal cases resolve as the product of decision

---

237 Marc L. Miller & Ronald F. Wright, *The Black Box*, 94 IOWA L. REV.125, 156 (2008).

238 *See* Bibas, *supra* note 33, at 987.

239 MEASURES FOR JUST., *Overview* (2018), https://measuresforjustice.org (last visited Sep. 3, 2019).

240 INTER-U. CONSORTIUM POL. & SOC. RES., *Uniform Crime Reporting Program Data Series*, https://www.icpsr.umich.edu (last visited Sep. 3, 2019).

241 U.S. SENT'G COMM'N, *About the Commission*, https://www.ussc.gov (last visited Sept. 3, 2019).

242 Besiki L. Kutateladze et al., *Prosecutorial Attitudes, Perspectives, and Priorities: Insights from the Inside*, MACARTHUR FOUND. 1, 1–2 (Dec. 2018), https://caj.fiu.edu/news/2018/prosecutorial-attitudes-perspectives-and-priorities-insights-from-the-inside/report-1.pdf; *see also* Andrew Pantazi, *What Makes a Good Prosecutor? A New Study of Melissa Nelson's Office Hopes to Find Out*, FLA. TIMES UNION (Mar. 9, 2018), https://www.jacksonville.com/news/20180309/what-makes-good-prosecutor-new-study-of-melissa-nelsons-office-hopes-to-find-out.

243 Enrico Francesconi, *On the Future of Legal Publishing Services in the Semantic Web*, FUTURE INTERNET, June 2018, at 1, 6.

244 Bruna Armonas Colombo et al., *Challenges When Using Jurimetrics in Brazil—A Survey of Courts*, FUTURE INTERNET, Oct. 2017, at 1.

making during this period. The existing literature provides anecdotal, experimental, and real-world data and information that makes a strong case for the existence of racial bias in prosecutors, criminal defense attorneys, and judges. Moreover, there is reason to believe that defendants' attitudes towards the justice system and its actors, as well as defendants' responses to perceived discrimination, are likely to lead to suboptimal decision making and a vicious cycle of compounding racial disparity. While a number of interventions have been tried and even more have been proposed, these have shown few signs of being efficacious. We endorse a novel way forward. Prosecutors should be provided with race-neutral baselines—anchors to their decision making as a case progresses—and prosecutors' offices should organize their institutional data and allow either internal or external parties to analyze it and produce regular and open reports. In addition to resulting in greater fairness in PAPT decision making, this will increase trust in the U.S. criminal justice system.