Big Data, Machine Judges, and the Legitimacy of the Criminal Justice System

Ric Simmons[*]

Predictive algorithms are rapidly spreading throughout the criminal justice system. They are used to more efficiently allocate police resources, identify potentially dangerous individuals, and advise judges at bail hearings and sentencing determinations. These algorithms have the potential to increase the accuracy, efficiency, and fairness of the criminal justice system. However, some criticize them on the grounds that they may reinforce pre-existing biases against minorities. But one aspect of these tools that has not yet been discussed in the literature is whether they will be accepted as legitimate. For centuries, human beings made these critical decisions that affect people's safety and liberty. But now, for the first time in human history, we are delegating large aspects of these decisions to machines. This Article addresses whether people will be willing to accept this change, and if not, how we can adapt the algorithms in order to make them more acceptable.

In determining whether criminal defendants are likely to accept predictive algorithms, the Article draws on the field of procedural justice, which sets out numerous factors that determine whether a participant in a judicial proceeding believes that the process is fair. The Article finds that predictive algorithms do not fare particularly well on these factors; they may not be seen as trustworthy or neutral, and they do not give defendants a significant opportunity to participate in the process. The Article suggests that criminal defendants would be more likely to view predictive algorithms as legitimate if the algorithms were made more transparent,

[*] Copyright © 2018 Ric Simmons. Chief Justice Thomas J. Moyer Professor for the Administration of Justice and Rule of Law, Moritz College of Law, The Ohio State University. I would like to thank Angela Lloyd, Dennis Hirsch, and the participants in the Big Data and Criminal Law Workshop at Moritz for their very helpful comments. I would also like to thank Kelsey Kornblut and Michael Wilhelm for their research assistance.

and if they were designed in a manner that does not use data tainted by past discriminatory practices.

The Article then examines whether the general population is likely to perceive predictive algorithms as legitimate. It examines various psychological barriers that people have with regard to accepting predictive algorithms. The Article presents an original empirical study of six hundred individuals who were presented with a hypothetical case in which a judge uses a predictive algorithm to assist in a bail hearing. The study indicates that individuals are likely to accept predictive algorithms as long as certain criteria are met.

TABLE OF CONTENTS

INTRODUCTION

Computer-run predictive algorithms[1] are now an integral part of our society.[2] We use them to match partners on online dating sites; recommend books and movies based on past preferences; and determine the best route to a certain destination.[3] More fundamentally, algorithms drive all internet search engines; coordinate many financial transactions; and help factories operate.[4]

The criminal justice system is not immune to this trend. Computer algorithms are currently used to provide more effective allocation of police resources;[5] to notify police of potentially dangerous individuals at specific locations;[6] to identify potential criminals based on their social media posts;[7] to guide efforts to intervene with individuals before they engage in criminal activity;[8] to advise judges making decisions about pretrial detention;[9] and to provide guidance to judges

[1] This paper will use the term "predictive algorithms" to refer to computer programs that use algorithms to make predictions and recommendations. Technically the term "algorithm" refers to any process or set of rules used to solve a problem, and so humans can use algorithms as well. Although this paper is focusing on the use of computer-run predictive algorithms in the criminal justice system, for brevity's sake it will use the shorthand "predictive algorithm" to refer to algorithms run by computers.

[2] One commentator noted that "[i]f every algorithm suddenly stopped working, it would be the end of the world as we know it." PEDRO DOMINGOS, THE MASTER ALGORITHM 1 (2015).

[3] Lee Rainie & Janna Anderson, *Code-Dependent: Pros and Cons of the Algorithm Age*, PEW RES. CTR. (Feb. 8, 2017), http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/.

[4] *See id.*; *see also* Louis Columbus, *10 Ways Machine Learning Is Revolutionizing Manufacturing*, FORBES (June 26, 2016, 8:17 PM), https://www.forbes.com/sites/louiscolumbus/2016/06/26/10-ways-machine-learning-is-revolutionizing-manufacturing/#713a113428c2.

[5] *See* Maurice Chammah, *Policing the Future*, MARSHALL PROJECT (Feb. 3, 2016, 7:15 AM), https://www.themarshallproject.org/2016/02/03/policing-the-future.

[6] Justin Jouvenal, *The New Way Police Are Surveilling You: Calculating Your Threat "Score,"* WASH. POST (Jan. 10, 2016), https://www.washingtonpost.com/local/public-safety/the-new-way-police-are-surveilling-you-calculating-your-threat-score/2016/01/10/e42bccac-8e15-11e5-baf4-bdf37355da0c_story.html.

[7] Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 HARV. L. & POL'Y REV. 15, 24 (2016).

[8] Matt Stroud, *The Minority Report: Chicago's New Police Computer Predicts Crimes, but Is It Racist?*, VERGE (Feb. 19, 2014, 9:31 AM), http://www.theverge.com/2014/2/19/5419854/the-minority-report-this-computer-predicts-crime-but-is-it-racist.

[9] Shaila Dewan, *Judges Replacing Conjecture with Formula for Bail*, N.Y. TIMES (June 26, 2015), https://www.nytimes.com/2015/06/27/us/turning-the-granting-of-bail-into-a-science.html.

at sentencing.[10] Soon they will likely be assisting police and magistrates in making reasonable suspicion and probable cause determinations. Like predictive algorithms in other fields, these tools show great promise in increasing the accuracy and efficiency of the criminal justice system. Nevertheless, they remain controversial, as they also have the potential to reinforce long-standing discriminatory practices by police and judges.[11]

This paper argues that on balance, adopting predictive algorithms will be a positive development for the criminal justice system. While many commentators argue that predictive algorithms pose a severe threat to the fairness of the criminal justice system,[12] these tools will increase the accuracy, efficiency, and fairness of many aspects of policing and adjudication if instituted properly. However, there is one aspect of predictive algorithms that poses a potentially fatal challenge to their widespread adoption: the possibility that no matter how accurate and fair they are, they will still not be accepted as legitimate methods of decision-making in the criminal justice system.

Predictive algorithms have been accepted as legitimate in many other aspects of our life, but most of those contexts are very different from their use in the criminal justice system. In some cases, such as when they route our financial transactions, predictive algorithms are invisible to the vast majority of people who are affected by them. In other cases, such as when they recommend movies, they are seen as relatively frivolous. But predictive algorithms in the criminal justice system are quite visible, in that judges openly refer to them when making bail and sentencing decisions, and these algorithms make decisions that affect people's dignity and liberty. It may be hard for people to accept the concept that computers are replacing humans in making these critical decisions.

Ordinary people have shown reluctance to embrace predictive algorithms that make significant decisions in other contexts. Consider the case of self-driving cars, which are driven by electronic sensors and algorithms instead of by human beings. These vehicles have the

[10] *See* Christopher Slobogin, *Risk Assessment*, in THE OXFORD HANDBOOK OF SENTENCING AND CORRECTIONS 196, 203-05 (Joan Petersilia & Kevin R. Reitz eds., 2012).

[11] *See* Jason Tashea, *Calculating Crime: Attorneys Are Challenging the Use of Algorithms to Help Determine Bail, Sentencing and Parole Decisions*, ABA J., Mar. 2017, at 54, 56, 58 [hereinafter *Calculating Crime*] (discussing the ongoing national debate about the use of algorithms in bail, sentencing, and parole decisions); *see also infra* Part I.B.

[12] *See infra* notes 25–36 and accompanying text.

potential to revolutionize travel: they will be safer; reduce traffic congestion; offer increased fuel efficiency; and make automobiles accessible to segments of the population that were unable to drive cars in the past.[13] Yet, many people are reluctant to adopt this new technology. A recent survey showed that only thirty to forty percent of Americans would definitely consider riding in a self-driving car, and fewer than twenty percent of Americans would definitely consider doing so in a heavily congested or high speed area.[14] Regardless of how many benefits self-driving cars offer, many will be reluctant to relinquish control of this traditionally human function to a machine.

This same reluctance to cede control to predictive algorithms is playing out in policing and courtrooms. For centuries, decisions in our criminal justice system have been made by individuals: police officers, judges, bureaucrats, and legislators. We are accustomed to these decision-makers and, for the most part, accept their legitimacy. For reasons which may be hard to describe, we are apprehensive about turning these decisions over to machines. This apprehension persists even if data scientists demonstrate that machines are more accurate in their predictions. And it persists even though we are well aware that the most well-intentioned human decision-makers are affected by implicit bias.

In other words, even if we can prove that predictive algorithms in the criminal justice system can produce decisions that are more fair, efficient, and accurate than human judgment, we risk delegitimizing the system if defendants subject to the decisions of the algorithms see them as illegitimate. Furthermore, these algorithms will never achieve widespread adoption in the criminal justice system unless the general public, too, views them as legitimate. This Article examines the various psychological barriers that people have regarding the use of predictive algorithms in the criminal justice system, and offers suggestions as to how to overcome these barriers.

---

[13] DANIEL HOWARD & DANIELLE DAI, PUBLIC PERCEPTIONS OF SELF-DRIVING CARS: THE CASE OF BERKELEY, CALIFORNIA 4-5 (2013), https://www.ocf.berkeley.edu/~djhoward/reports/Report%20-%20Public%20Perceptions%20of%20Self%20Driving%20Cars.pdf.

[14] These are the results of a State Farm/Bloomberg Government survey from September 16, 2016. STATE FARM & BLOOMBERG GOV'T, PUBLIC PERCEPTIONS OF DRIVERLESS CARS 2 (2016), https://newsroom.statefarm.com/download/271091/bloombergstatefarmsurveyreport-final.pdf. Other surveys show similar results. For example, a 2017 Pew Survey showed that fifty-six percent of respondents would not ride in an autonomous vehicle. AARON SMITH & MONICA ANDERSON, PEW RESEARCH CTR., AUTOMATION IN EVERYDAY LIFE 31 (2017), http://assets.pewresearch.org/wp-content/uploads/sites/14/2017/10/03151500/PI_2017.10.04_Automation_FINAL.pdf.

Part I of this Article provides a brief overview of the potential for predictive algorithms to increase the accuracy and efficiency of the criminal justice system and then discusses the concerns that have been raised about these tools, including the problem of biased data and the argument that it is fundamentally unfair to judge individuals based on broad group characteristics.[15] Part II focuses on how criminal defendants perceive predictive algorithms, and uses the principles of procedural justice to examine whether criminal defendants will find these algorithms to be fair. Part III explores more generally how the general public is likely to react to the widespread use of these tools, and presents the results of a survey which suggests that people are comfortable with machines making these decisions. The Article then concludes with some suggestions for how to regulate and use predictive algorithms to increase their acceptance among defendants and the general public.

I. THE RISE OF BIG DATA

A. The Promise of Predictive Algorithms

Across disciplines, predictive algorithms consistently outperform human judgments. Even before the dramatic increase in data collection and data processing known as "big data," evidence-based algorithms outperformed human judgment in areas such as predicting academic performance and diagnosing medical conditions.[16] A meta-analysis in the year 2000 of 136 studies showed that "algorithms outperformed human forecasters by ten percent on average and that it was far more common for algorithms to outperform human judges than the opposite."[17]

Given this success, it is no surprise that police departments and courts are incorporating predictive algorithms in crime prevention, bail hearings, and sentencing hearings. Jurisdictions in nearly all fifty states have begun to use these tools.[18] As these models begin to access

[15] As we will see, "broad group characteristics" refers not just to traditional suspect classes such as race and religion, but also otherwise innocuous characteristics, such as employment status or age.

[16] Berkeley J. Dietvorst, Joseph P. Simmons & Cade Massey, *Algorithmic Aversion: People Erroneously Avoid Algorithms After Seeing Them Err*, J. EXPERIMENTAL PSYCHOL. 114, 114 (2014) (citing studies from before 1954 in predicting academic performance and parole violations, and studies from before 1979 and 1989 for clinical diagnosis and forecasting graduate students' success).

[17] *Id.*

[18] *Algorithms in the Criminal Justice System*, ELEC. PRIVACY INFO. CTR.,

and process the massive amounts of information available from the big data revolution, we see increased evidence that a computer's predictive algorithms make more accurate predictions than a judge's clinical judgments. For example, after the state of Kentucky adopted a risk assessment tool for pretrial detention determinations that had been developed by the Arnold Foundation, more defendants were released pending trial, yet the average arrest rate for released defendants fell significantly.[19] One year later, the courts in Lucas County, Ohio adopted the same tool and found an even greater improvement: the percentage of defendants released without bail increased from fourteen percent to nearly twenty-eight percent, while the percentage of defendants who committed crimes after being released dropped from twenty percent to ten percent.[20] A Washington Post study of 5000 defendants in Florida found that the risk scores of the COMPAS predictive algorithm developed by the Northpointe corporation[21] were "highly predictive" of reoffending, in that defendants assigned the highest risk score reoffended eighty-one percent of the time, and those assigned the lowest risk score reoffended at a twenty-two percent rate.[22]

https://epic.org/algorithmic-transparency/crim-justice/ (last visited Aug. 30, 2018). In some cases, their use is mandated by statute or court rule. In other jurisdictions, their use is recommended but not required. The vast majority of jurisdictions use one of three systems: the Correctional Offender Management Profiling for Alternative Sanctions ("COMPAS"), which is the product of a private company called Northpointe; Public Safety Assessment ("PSA"), which was created by a private company called Multi-Health Systems, and Level of Service Inventory Revised ("LSI-R"), which was created by the Laura and John Arnold Foundation. *Id.*; *see also* NAT'L CTR. FOR STATE COURT, APPENDIX A PROFILES OF ASSESSMENT INSTRUMENTS 2-13 (last visited Aug. 30, 2018), http://www.ncsc.org/~/media/Files/PDF/Services%20and%20Experts/Areas%20of%20expertise/Sentencing%20Probation/RAN%20Appendix%20A.ashx (describing seven different state-specific risk assessment tools).

[19] *See* Jason Tashea, *Bond Ratings: Kentucky Is Testing a New Assessment Tool to Determine Whether to Keep Defendants Behind Bars*, ABA J., April 2015, at 15, 16 [hereinafter *Bond Ratings*] (explaining that the percentage of defendants who were released increased from 68% to 70% while the average arrest rate declined from 10% to 8.5%). The Arnold Foundation developed its new PSI tool by analyzing a data set of 1.5 million cases. *Id.*

[20] Tashea, *Calculating Crime*, *supra* note 11, at 58.

[21] COMPAS stands for Correctional Offender Management Profiling for Alternative Sanctions.

[22] Sam Corbett-Davies et al., *A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It's Actually Not That Clear*, WASH. POST (Oct. 17, 2016), https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/?noredirect=on&utm_term=.ea069faac314. The *Washington Post* study was re-analyzing data from a *ProPublica* study which claimed racial bias in the COMPAS algorithm. *Id.; see also* Julia Angwin et al.,

Big data algorithms promise not just greater accuracy in their predictions, but greater efficiency. When the Arnold Foundation reviewed Kentucky's pre-existing method for pretrial detention determinations, it noted that much of the information that the court was collecting did not help to predict the defendant's future behavior.[23] The Foundation's new risk-assessment tool lowered the number of risk factors from dozens of factors down to nine, which lowered the average length of the pre-arraignment interview between defendants and court personnel from twenty minutes to three.[24]

B. Concerns About Predictive Algorithms

Of course, accuracy is not the only goal of the criminal justice system. Critics of the use of predictive algorithms in criminal contexts point to a number of potential problems. The most significant problem is that the data used by these algorithms may already be tainted by decades of discriminatory practices in the criminal justice system. Although no criminal justice predictive algorithm will directly consider race as a factor in making a decision, the use of predictive algorithms still has the effect of confirming pre-existing biases in policing and judging. As Professor Andrew Ferguson notes, the use of predictive algorithms in policing can result in a vicious cycle, since "many of the variables [considered by the algorithm] directly correlate with racially discriminatory law enforcement practices."[25] Racial minorities have disproportionate contact with the criminal justice system: for example, although whites and African-Americans use marijuana at the same rates, African-Americans are 3.73 times more likely than whites to be arrested for marijuana possession.[26] Since many predictive algorithms rely upon prior arrests or convictions as a factor in predicting future criminal activity, predictive algorithms will perpetuate this racial disparity unless efforts are made to remedy this

Machine Bias, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. COMPAS gathers 137 different responses from the defendant and combines that data with the defendant's prior criminal record to estimate the defendant's chances of committing a future crime. *Id.*

[23] *See* Tashea, *Bond Ratings*, *supra* note 19, at 16. For example, the defendant's personal references and employment history had no predictive value. *Id.*

[24] *Id.* The Arnold Foundation's goal was to determine "the least amount of information that could still determine a released defendant's failure to appear or likelihood of committing a new crime." *Id.*

[25] ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT 47 (2017).

[26] *Id.* at 48.

problem. Even if predictive algorithms do not focus upon prior criminal activity, they often use other factors which are highly correlated to race, including socio-economic status or neighborhood of residence, leading to results which disproportionately harm minorities.[27]

But the racially disproportionate raw data is not a problem confined to big data's predictive algorithms. Human police officers and judges use the same data when making their own predictions of where to patrol or whether to set bail. The same vicious cycle applies when humans make the decisions: historical crime data leads police officers to areas where crime is most likely to occur; the police disproportionately patrol in those neighborhoods (which tend to be poor and non-white) and make more arrests; and those arrests are then used by police in the future to continue to disproportionately patrol those neighborhoods.

Critics of predictive algorithms respond that when machines use this biased data and make a prediction, it carries a stronger implication of neutrality. As one critic of predictive algorithms writes, these algorithms create an illusory "technocratic framing" of who is dangerous and who deserves greater punishment, even though the algorithms' conclusions are based on the same flawed data.[28] On the other hand, predictive algorithms give us an opportunity to overcome bias because we can monitor the data that the algorithms use and, if need be, correct for pre-existing racial discrimination. For example, if the data show that African-Americans are arrested at disproportionate rates for possessing marijuana, the program could be designed to discount marijuana arrests of Africa-Americans by the appropriate ratio.[29] In fact, big data tools allow us to identify and correct these pre-existing biases with greater precision than before. In the past, police and judges may have been only dimly aware of the racial disparities in the factors they were using to predict future criminal activity, and those that were aware of those disparities had limited options in remediating the problem. Today, big data can measure the pre-existing racial biases, and properly designed predictive algorithms can be programmed to avoid those factors or adjust the data to counteract its discriminatory effect.[30]

---

[27] *See* Angwin, *supra* note 22.

[28] Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803, 806 (2014) [hereinafter *Evidence-Based*].

[29] *See infra* Part II.B. for further discussion of this issue.

[30] *Cf.* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 727-28 (2016). Given the amount of information now known about the

Finally, some critics argue that even if predictive algorithms could be designed to have no disproportionate effect on racial minorities, it is fundamentally unfair to base these predictions on immutable characteristics, such as age or gender, or indeed any broad group characteristic, such as employment status or neighborhood of residence.[31] Using status as a factor in determining a person's sentence contravenes standard punishment theory, since it makes punishment dependent upon who a person is rather than what he has done.[32] Although human police officers and judges base their decisions on a defendant's status or broad group characteristics, many commentators argue that it is wrong to formally codify these factors in a computer program.[33] This is known as the "individualized suspicion" problem: even if predictive algorithms could lead to more accurate results, it would be legally and morally wrong to take action against a person based on membership in a specific class or other group status.[34] This is why it is critical for predictive algorithms to also consider the defendant's specific conduct as a factor in its predictions; they must be designed so that they do not make predictions based solely on status. In other words, the prediction cannot be based only on who the person is; it must also be based in part on what the person does.[35]

A full debate regarding the potential benefits and challenges presented by predictive algorithms is well beyond the scope of this

exact extent of biasing in policing, an algorithm could estimate the over-representation of minorities in the existing data of police stops and arrests and adjust the data accordingly before performing its calculations.

[31] Devlin Barrett, *Holder Cautions on Risk of Bias in Big Data Use in Criminal Justice*, WALL ST. J. (Aug. 1, 2014), https://www.wsj.com/articles/u-s-attorney-general-cautions-on-risk-of-bias-in-big-data-use-in-criminal-justice-1406916606.

[32] Starr, *Evidence-Based*, *supra* note 28, at 817 (discussing Bernard Harcourt's theory in BERNARD HARCOURT, AGAINST PREDICTION PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE 173-92 (2007)).

[33] *See, e.g.*, Sonja B. Starr, *Sentencing, by the Numbers*, N.Y. TIMES (Aug. 10, 2014), https://www.nytimes.com/2014/08/11/opinion/sentencing-by-the-numbers.html?ref=todayspaper&_r=1 [hereinafter *Sentencing, by the Numbers*]. We will consider this argument in more detail in Part III.A.3.

[34] *See* United States v. Cortez, 449 U.S. 411, 418 (1981) ("[T]he process . . . must raise a suspicion that the particular individual being stopped is engaged in wrongdoing."). *See generally* BERNARD E. HARCOURT, AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE 173-92 (2007). The Supreme Court has held that it was illegal to stop and frisk an individual based only on the individual's presence in a tavern where drugs were being sold, ruling that there must be reasonable suspicion "directed at the person to be frisked." Ybarra v. Illinois, 444 U.S. 85, 94 (1979).

[35] *See* Andrew Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327, 388 (2015).

paper. In this brief section, we have seen that predictive algorithms have the potential to revolutionize the criminal justice system, but that they must minimize the racially discriminatory impact of the data they use. This will require greater transparency so that police departments, judges, and policymakers can evaluate whether the algorithms meet these criteria.[36]

But even after predictive algorithms overcome these problems, they will face one final challenge: whether they will be accepted as legitimate decision-makers in the criminal justice system, or whether their adoption will undermine the system's legitimacy. If supporters of predictive algorithms want to see widespread adoption of these programs by courts and law enforcement agencies, they must do more than merely improve their fairness and transparency. They must also convince the general population — and the defendants impacted by these algorithms — to accept them as legitimate. The next section will examine how criminal defendants perceive the use of predictive algorithms in their own cases. The following section will examine how the general population is likely to view the use of these algorithms throughout the criminal justice system.

II. PROCEDURAL JUSTICE

Procedural justice is the field of study concerned with the perceived fairness of legal systems. It applies psychology to the law to determine what aspects of our legal system lead people to believe the system is legitimate. Procedural justice focuses on the procedural aspects of dispute resolution, not the substantive outcomes; and it is concerned with perceived fairness, not actual fairness.[37] Numerous psychological

[36] *See* AARON RIEKE, DAVID ROBINSON & HARLAN YU, CIVIL RIGHTS, BIG DATA, AND OUR ALGORITHMIC FUTURE 20 (2014) (discussing how transparency in police surveillance technologies is vital to mitigate discriminatory police practices); EXEC.OFFICE OF THE PRESIDENT, BIG DATA: A REPORT ON ALGORITHMIC SYSTEMS, OPPORTUNITY, AND CIVIL RIGHTS 21 (2016) (discussing how transparency and accountability on data input and processes can help ensure that data and algorithmic systems are not "used in ways that exacerbate unwarranted disparities in the criminal justice system"); Joh, *supra* note 7, at 40 (discussing the need for transparency and accountability measures to address concerns raised by expanded surveillance discretion); Laurel Eckhouse, *Big Data May Be Reinforcing Racial Bias in the Criminal Justice System*, WASH. POST (Feb. 10, 2017), https://www.washingtonpost.com/opinions/big-data-may-be-reinforcing-racial-bias-in-the-criminal-justice-system/2017/02/10/d63de518-ee3a-11e6-9973-c5efb7ccfb0d_story.html?utm_term=.dcda87fcce40 (urging for more transparency of risk assessment tools to determine whether they have disparate impacts on defendants of different races).

[37] *See* Paul G. Chevigny, *Fairness and Participation*, 64 N.Y.U. L. REV. 1211, 1211

studies show that whether an individual believes a particular interaction with authority is fair and legitimate depends primarily on whether they think the procedure is fair, and does not depend upon the ultimate outcome of the interaction.[38] In a complex society, individuals regularly receive and are denied a variety of benefits. They are similarly subjected to a variety of costs. Because it is challenging to keep track of each of these substantive outcomes, individuals usually focus on the procedure used to assess their own costs and benefits in deciding whether or not a particular interaction is fair.[39] Furthermore, individuals in a complex society often disagree about the appropriate distribution of costs and benefits but are more likely to agree on whether a certain procedure is fair.[40]

This section will use the principles of procedural justice to determine whether the subjects of the criminal justice system — the criminal defendants — believe the use of predictive algorithms is fair and legitimate.[41] Obviously criminal defendants represent a diverse group of people, and they may not all react the same way to the use of predictive algorithms in their cases. However, since they are the ones most directly affected by the use of these algorithms, their reactions are especially significant. Although defendants, by definition, participate in the criminal justice system involuntarily, it is nevertheless crucial that they believe the criminal justice system treats them more or less fairly. If criminal defendants believe that a

(1989).

[38] *See, e.g.*, TOM R. TYLER, WHY PEOPLE OBEY THE LAW 73-74 (1990) [hereinafter OBEY THE LAW]. Citing six separate psychological studies, Professor Tyler notes that:

> Recent research confirms that people evaluate their experience in procedural terms. Such procedural effects have been found in trials as well as in other procedures used to resolve disputes, including plea bargaining, mediation, and decision making by police officers . . . . Wherever procedural issues have been studied they have emerged as an important concern to those affected by the decisions.

Id. at 74 (citations omitted).

[39] *Id.*

[40] *Id.*; *see also* Tom Tyler et al., *Procedural Justice in Felony Cases*, 22 LAW & SOC'Y REV. 483 (1988) (Interviews with over 400 prisoners showed that procedural justice, rather than length of sentence, was a substantial factor in determining whether the prisoners were satisfied with their sentence.).

[41] Of course, it is important for every member of society, not just criminal defendants, to perceive that the criminal justice system is legitimate. This section focuses only on how criminal defendants are likely to view the use of predictive algorithms in the criminal justice system. See *infra* Part III for a discussion and an attempt to measure how the general public views that phenomenon.

particular aspect of the criminal justice is unfair, that belief delegitimizes and eventually destabilizes the entire criminal justice system.[42] Although *actual* fairness is important to ensure the system is moral and just, perceived fairness is critical to ensure that the system is accepted, or at least tolerated, by those being judged. Perceived fairness encourages compliance with the law; in other words, people obey the law primarily because they believe that it is fair and legitimate, not because they fear sanctions if they transgress.[43] A legal system with high levels of procedural justice is more stable.[44]

On a more practical level, a well-ordered society cannot exist based on the threat of sanction alone; it must primarily depend on voluntary compliance, which depends, in part, on the perceived legitimacy of the system. At a time when many institutions of the criminal justice system are under attack, enhancing the procedural justice of the system is critical, as it builds confidence in institutions, thereby enhancing respect for the police and the courts.[45]

---

[42] There is always the danger that authorities will be able to provide a procedure that appears fair but in fact is not, thus creating a "false consciousness": an inaccurate sense of legitimacy on the part of the people that allows authorities to abuse their position and impose procedures, and distributive outcomes, which are "objectively" unfair. TYLER, OBEY THE LAW, *supra* note 38, at 111.

[43] *Id.* at 60-62. Professor Tyler interviewed over 1,500 people to determine why they obeyed a wide variety of criminal laws and determined that belief in the legitimacy of the legal system was a significant influence in their decision not to break the law. *See* Margaret B. Kwoka, *Leaking and Legitimacy*, 48 UC DAVIS L. REV. 1387, 1421-22 (2015) (summarizing studies that link the belief in the procedural justice of the laws to the likelihood of an individual breaking those laws in the context of immigration violations and tax fraud).

[44] TYLER, OBEY THE LAW, *supra* note 38, at 107 ("[F]air procedures can act as a cushion of support when authorities are delivering unfavorable outcomes. If unfavorable outcomes are delivered through procedures viewed as fair, the unfavorable outcomes do not harm the legitimacy of legal authorities."). For example, studies have shown that defendants who plea bargained cases found the process to be a fairer procedure than resolving the criminal issues at trial, in part because the defendants had more control over the process of a plea bargaining procedure. *See* Michael O'Hear, *Plea Bargaining and Procedural Justice*, 42 GA. L. REV. 407, 412 (2008) (arguing for greater procedural justice in the plea bargaining system in order to increase defendants' satisfaction with the outcome of the plea bargains and to enhance their compliance with legal rules generally); Tom R. Tyler, *Citizen Discontent with Legal Procedures: A Social Science Perspective on Civil Procedure Reform*, 45 AM. J. COMP. L. 871 *passim* (1997) (showing a crisis of public confidence in the American justice system resulting from people's dissatisfying personal experiences with the court and the law).

[45] *See, e.g.*, Rebecca Hollander-Blumoff, *The Psychology of Procedural Justice in the Federal Courts*, 63 HASTINGS L.J. 127, 129 (2011) (examining the procedural justice aspects of our civil federal court system); Hon. William G. Young & Jordan M. Singer,

Finally, criminal defendants must, to some extent, acquiesce in the criminal justice system. From voluntary interactions with pretrial services officers, corrections officers, and probation officers to plea bargaining and cooperation with their attorneys throughout the process, "the normally coercive machinery of criminal justice depends to a surprising extent on the cooperation of the defendants moving through the system."[46] When choosing between two potential procedural alternatives in the criminal justice system, such as setting bail based either solely on the clinical judgment of a judge or in part on the risk factors assessed by an algorithm, the degree to which defendants believe a given procedure to be fair is a relevant factor in deciding which practice to use.

Procedural justice scholars have identified four factors that determine whether an individual believes a given procedure is fair: interpersonal respect, neutrality, trustworthiness, and process control.[47] The following sections will consider how predictive algorithms affect each of these four factors.

A. Interpersonal Respect

The first factor used to determine whether an individual will find a process fair is whether the process treats individuals with dignity and respect.[48] This includes interacting with police officers during *Terry* stops, arrests, and booking, as well as interacting with judges in the courtroom.

This factor is unlikely to matter in evaluating the perceived legitimacy of big data algorithms, since the individuals who are subject to the algorithm's decisions will not be interacting directly with the algorithm or its designer. Instead, the individuals will interact with

Bench Presence: Toward a More Complete Model of Federal District Court Productivity, 118 PENN ST. L. REV. 55, 74 (2013) ("Procedural fairness . . . is seen as a necessary value both for generating fairer outcomes and for building public confidence in the judicial system's ability to generate those outcomes.").

[46] Michael O'Hear, *Explaining Sentences*, 36 FLA. ST. U. L. REV. 459, 478 (2009) [hereinafter *Explaining Sentences*].

[47] Tom R. Tyler, *Social Justice: Outcome and Procedure*, 35 INT'L J. PSYCHOL. 117, 121 (2000) [hereinafter *Social Justice*]; Tom R. Tyler & Hulda Thorisdottir, *A Psychological Perspective on Compensation for Harm: Examining the September 11th Victim Compensation Fund*, 53 DEPAUL L. REV. 355, 380-82 (2003). I have written about the application of procedural justice to predictive algorithms elsewhere. *See* Ric Simmons, *Big Data and Procedural Justice: Legitimizing Algorithms in the Criminal Justice System*, 15 OHIO ST. J. CRIM. L. 573 (2018). Some parts of the following analysis have been adapted from this earlier essay.

[48] *See* Tyler, *Social Justice*, *supra* note 47, at 122.

human police officers, prosecutors, or judges who are acting on the software's recommendation. Regardless of whether these human actors use predictive algorithms to guide their decisions, they still choose whether to treat defendants with dignity and courtesy.

There is a danger that police and judges who rely on predictive algorithms will perceive criminal defendants as merely a collection of numbers and probabilities rather than human beings, and as a result treat them more like objects on an assembly line rather than as individual people. Furthermore, the very use of computer algorithms to advise police and judges in these decisions may affect the perceived dignity of defendants who are subject to these decisions. Individuals may feel that their value as a human being is lessened when they are frisked on the street or held in custody without bail because a computer said it was the proper course of action. One way to alleviate this potential problem is for police and judges to explain the reasoning behind their actions to the defendants. If the decision that is being made is based in part on a predictive algorithm, the authorities should explain the details of the algorithm to the defendants rather than just saying that a machine mandated the decision. This will require both that the algorithms themselves are transparent so that the police and judges understand the factors that led to the algorithm's decision, and extra efforts on the part of the authorities to explain these factors to the defendants. As we will see below, this extra effort will also help in establishing trustworthiness, another factor that is critical to procedural justice.

B. Neutrality

Neutrality refers to whether a decision-maker follows rules impartially and makes objective, unbiased decisions grounded in facts.[49] Predictive algorithms hold great promise for increasing both actual and perceived neutrality of the decision-makers, because the algorithm follows the same rules in the same way for every decision it makes.

As discussed in Part I, however, predictive algorithms currently face great challenges in this area, since often they reinforce pre-existing biases and partialities built into the system.[50] Particular groups in society are more likely to be convicted because of their race or the

[49] *Id.* This was first thought to be the most important aspect of a procedure in creating legitimacy, but studies have shown that it ranks behind process control and trustworthiness. *Id.*

[50] *See supra* notes 25–30 and accompanying text.

neighborhood in which they live. Therefore, a sentencing algorithm that uses, as key factors, prior convictions or home address to determine the length of the sentence will exacerbate existing inequalities. Even if an algorithm avoids using race or proxies for race as factors in determining risk, its results may still have a disproportionate impact on black and Latino defendants. Indeed, the most common critique of these predictive algorithms is that they predict black defendants to have a higher recidivism rate than white defendants. A 2016 ProPublica study analyzed the risk scores of thousands of defendants that the COMPAS algorithm evaluated and concluded that the false positive rate was twice as high for black defendants as it was for white defendants.[51]

Part of the problem in evaluating the neutrality of predictive algorithms is that different people may have different definitions of what is "neutral." Northpointe, which designed the COMPAS algorithm, points out that amongst defendants who have identical risk scores, the actual recidivism rate is identical for black defendants and white defendants.[52] This is strong evidence that the algorithm is "neutral," in that its results are identical across race. However, because the overall recidivism rate is higher for black defendants than white defendants, the program is statistically guaranteed to have a higher false positive rate for blacks than for whites. In other words, in order to be race neutral in its results, COMPAS — or any predictive system — will have to have more false positives for blacks than whites.[53]

[51] Angwin, *supra* note 22. A false positive occurs when the algorithm examines all the data at arraignment and then predicts that a defendant will re-offend, but in fact the defendant does not re-offend. A false negative, in contrast, occurs when the algorithm predicts that the defendant will not re-offend but in fact the defendant does re-offend.

[52] *See* Corbett-Davies et al., *supra* note 22.

[53] *See id.* The data collected by *ProPublica*, for example, contained 5000 defendants who were assigned COMPAS scores in Broward County, Florida. Among those 5000, the algorithm predicted the recidivism rate of blacks and whites with equal accuracy — for example, a white defendant who received a risk score of seven had a recidivism rate of sixty percent, while a black defendant who received the same risk score had a recidivism rate of sixty-one percent. However, the overall recidivism rates were significantly higher for black defendants than for white defendants (fifty-two to thirty-nine percent). Thus, black defendants are more likely to be classified as medium or high risk than white defendants (fifty-eight to thirty-three percent), which will result in a higher percentage of black defendants than white defendants being incorrectly flagged as high-risk. As the authors note, "If the recidivism rate for white and black defendants is the same within each risk category, and if black defendants have a higher overall recidivism rate, then a greater share of black defendants will be classified as high risk. And if a greater share of black defendants are classified as high risk, then . . . a greater share of black defendants who do not reoffend will also be

It is hard to say how these two dueling and incompatible values — neutrality in results versus identical false positive rates — translates into the procedural justice realm. In theory, neutrality in results should be the most important factor — that is, a black defendant must believe, as is true in the case of COMPAS, that a given score is equally accurate for him as it would be for a white defendant. But the disparity in the false positive rate, combined with the fact that low income level or prior convictions (which are both correlated to race) increase a defendant's risk score, may understandably cause black and Latino defendants to perceive these algorithms as not "neutral" to their race.

Of course, these critiques apply as much to individual human judgments as they do to an algorithm. The racial disparities in our criminal justice system are well documented, and pre-date the use of predictive algorithms. As of 2014, for example, before computer algorithms began influencing sentencing, 2.7% of black makes and 1.1% of Latino males were serving sentences of at least one year in prison, while only .5% of white males were serving a similar sentence.[54] And the mathematical rule that mandates a higher false positive rate if the results are neutral and the underlying recidivism rates are unequal holds true regardless of whether a machine or a human is making the prediction.[55]

As noted in Part I,[56] the only way to achieve truly neutral decision-making is to remove the race-based criteria from the predictive algorithm or remove all implicit bias from the human decision-maker. The former seems easier than the latter, since algorithms could theoretically be programmed to discount any factors which have been tainted by prior racial discrimination. For example, assume that a study shows that black residents and white residents commit a certain type of crime with equal frequency, but because of inequities in the system, black residents are more likely to be convicted of the crime.[57] Since prior criminal activity is a significant factor in predicting whether a defendant will re-offend, a predictive algorithm is likely to

classified as high risk." *Id.*

[54] E. ANN CARSON, U.S. DEP'T OF JUSTICE, PRISONERS IN 2014, at 15 (2015), https://www.bjs.gov/content/pub/pdf/p14.pdf.

[55] *See supra* note 53 and accompanying text.

[56] *See supra* note 28 and accompanying text.

[57] There are certainly studies which show that police focus their attention and resources disproportionately on black communities, even when holding the crime rate constant, which would lead to the conclusion that more arrests would occur in predominantly black communities than in predominantly white communities. *See, e.g.*, Katherine Beckett et al., *Drug Use, Drug Possession Arrests, and the Question of Race: Lessons from Seattle*, 52 SOC. PROBS. 419 (2005).

use this factor in determining the risk of recidivism. Thus, if the unadjusted prior conviction numbers are used, the recidivism rates for black defendants will be inaccurately (and unfairly) higher than those for white defendants. Once this fact has been established, programmers can alter their algorithms to remove the pre-existing inequity by discounting the number of prior convictions of this crime by the appropriate amount for a black defendant, or by increasing the number of the prior convictions by the appropriate amount for a white defendant. Such adjustments could be made for all factors in which past racial bias has been documented, and they could be integrated into the decision-making process at every level, from determining probable cause to setting bail to determining an appropriate sentence. Even if we could trust every human judge to be aware of these examples of pre-existing racial bias, and even if we could be confident that each judge would be sympathetic to the need to adjust decisions appropriately, it would be impossible for any human being to take into account all of these potential discounts with the appropriate level of precision. Only computer algorithms could hope to do this effectively. Thus, predictive algorithms have at least the potential to achieve greater neutrality than human beings, if they are designed appropriately and if the defendants who are subject to their decisions are aware of these designs.

C. Trustworthiness

In the field of procedural justice, trustworthiness can be established in two ways. First, a participant can feel a personal connection to the decision-maker; for example, the participant may perceive that the decision-maker comes from the same group as the participant. Second, the decision-maker can thoroughly explain his or her decisions.[58] Big data algorithms will rarely achieve trustworthiness through a personal connection, since most defendants will not feel any kinship to a computer program.[59] Thus, defendants will have to understand the

[58] *See, e.g.*, Tom R. Tyler & E. A. Lind, *A Relational Model of Authority in Groups*, *in* 25 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY 115, 155-57 (Mark P. Zanna ed., 1992) (explaining the preconditions for the effective functioning of authorities).

[59] Of course, the level of kinship that defendants feel with a judge under a traditional system will vary greatly. A defendant is unlikely to feel a connection with a judge who comes from a different economic class, has a different level of education, or is of a different ethnicity. *See, e.g.*, PEW RESEARCH CTR., MULTIRACIAL IN AMERICA: PROUD, DIVERSE AND GROWING IN NUMBERS 64 (2015), www.pewsocialtrends.org/2015/06/11/multiracial-in-america/ (A survey finds that relatively few adults say they have a lot in common with those who do not share their own racial background.);

reasoning undergirding predictive algorithms in order to feel that the algorithm is trustworthy. As with the interpersonal respect factor, this factor is harder to achieve for predictive algorithms than it would be for human decision-makers. Psychologists developed a concept known as the "theory of mind," which explores the way that human beings can understand mental states and thus the decisions and behavior of others.[60] Most people will experience a decision by a police officer or a judge and understand, or at least believe they understand, the reasoning behind it. This helps make the decision seem comprehensible and thus legitimate.

A predictive algorithm receives no such benefit of the doubt; thus it will only appear trustworthy if the criminal defendant understands how the algorithm reached its decision. This will mean that the human decision-maker needs to explain the reasoning behind the algorithm to the defendant, which in turn means that the human decision-maker must be able to understand how the algorithm works. Thus, as with the interpersonal respect and neutrality factors, greater transparency in the algorithm is critical to ensuring that the defendant believes the process is fair.

D. Participation in the Process

The final factor is known as "process control," or the individual's belief that they can participate in the process and express their views to the decision-maker. Studies have shown that individuals value this ability to participate even if their participation has no ability to affect the outcome of the procedure.[61] For example, an individual will be

Alberto Alesina & Eliana La Ferrara, *Who Trusts Others?*, J. PUB. ECON. 207, 231 (2002) (Study finds that factors associated with low trust include belonging to a group that historically felt discriminated against, being economically unsuccessful in terms of income and education, and living in a racial mixed community and/or in one with a high degree of income disparity.); Sandra Susan Smith, *Race and Trust*, ANN. REV. SOC. 453, 456 (2010), http://sociology.berkeley.edu/sites/default/files/faculty/Smith/RACE%20AND%20TRUST.pdf (discussing how "members of minority [racial] groups report substantially more misanthropy (less trust) than members of the majority").

[60] *See* David Premack & Guy Woodruff, *Does the Chimpanzee Have a Theory of Mind?*, 1 BEHAV. & BRAIN SCI. 515, 515 (1978).

[61] *See, e.g.*, E. Allan Lind et al., *Voice, Control, and Procedural Justice: Instrumental and Noninstrumental Concerns in Fairness Judgments*, 59 J. PERSONALITY & SOC. PSYCHOL. 952 (1990). One study found that allowing victims to testify at pretrial conferences increased the victim's perception of the fairness of the process even if their arguments had no effect on the ultimate sentence given to the defendant. *See* Anne M. Heinz & Wayne A. Kerstetter, *Pretrial Settlement Conference: Evaluation of a Reform in Plea Bargaining*, 13 LAW & SOC'Y REV. 349, 363 (1979).

more likely to see a process as fair if they are able to tell their story directly to the decision-maker, regardless of whether the story changes the decision.[62]

Even without the use of predictive algorithms, the criminal justice system fares poorly in the category of process control, since a defendant is represented by counsel during all formal procedures and thus rarely gets a chance to speak to the decision-maker.[63] However, predictive algorithms exacerbate this problem. When defendants are stopped or arrested by a police officer, they often attempt to explain their side of the story; if the police are making their decision based at least in part on the advice of a predictive algorithm, the defendant will feel less ability to affect the outcome of the interaction. Similarly, one of the few times that defendants get a chance to speak directly to a judge is at sentencing, when they can address the judge in support of their plea for a lighter sentence.[64] In the traditional model, the defendant may gain a sense of process control by speaking to the judge, even if the defendant has been warned that the judge is unlikely to change the ultimate sentence in response to the arguments. If, in contrast, defendants know that the ultimate decision has effectively already been made by a computer program, they will know that they are, in fact, not speaking to the true decision-maker. Indeed, that the decision has already been made even before the sentencing hearing begins. This will be the exact opposite of process control: the defendant will have no chance for real participation in the process and no opportunity to present his case to the true decision-maker in the case.

E. Improving the Procedural Justice Effect of Predictive Algorithms

Given how predictive algorithms are currently designed and applied, they likely decrease the level of procedural justice that defendants experience in the criminal justice system for a number of reasons. First, they are likely to reduce the level of interpersonal respect that

---

[62] *See* Tom R. Tyler et al., *Influence of Voice on Satisfaction with Leaders: Exploring the Meaning of Process Control*, 48 J. PERSONALITY & SOC. PSYCHOL. 72, 79 (1985).

[63] *See, e.g.*, Alexandra Natapoff, *Speechless: The Silencing of Criminal Defendants*, 80 N.Y.U. L. REV. 1449, 1458 (2005) (explaining how represented criminal defendants rarely speak since they are encouraged to be quiet and to let their lawyers do the talking).

[64] *See id.* at 1464-65 (discussing how defendants are entitled to address the court at sentencing and how sentencing is the hearing at which the court is most likely to hear the defendant speak); *see also* FED. R. CRIM. P. 32(i)(4)(A)(ii) (mandating that the court allow the defendant to speak or present any information to mitigate the sentence before imposing the sentence).

authorities give to defendants. Second, they may be perceived as not being neutral if they continue to perpetuate racial disparities in the system. Finally, defendants are less likely to feel a personal connection with them or understand why they come to their decisions, and defendants will experience a loss of process control when subject to their decisions.

There are two solutions to these problems. The first is maintaining a level of human interaction between the authority and the defendant. A defendant who is merely told by the police officer or judge that a machine has made the decision to arrest him or hold him in on bail will likely experience low levels of interpersonal respect, trustworthiness, and process control. But a defendant who hears an explanation from the police officer or judge is likely to experience higher levels of interpersonal respect and trustworthiness, and a defendant who at least believes that the police officer or judge is listening to his side of the story will not feel deprived of process control.[65]

This type of human interaction will not be possible without the second solution: increasing the transparency of the algorithms themselves. We have already seen that transparency is substantively important to both ensure that predictive algorithms are not racially biased and to guarantee they include an element of individualized suspicion. This section has shown that describing the workings of the algorithms may increase the level of interpersonal respect that defendants feel, enhance the perceived neutrality of predictive algorithms, and make them seem more trustworthy. Thus, greater transparency for these algorithms is critical both for substantive fairness and procedural justice. Unfortunately, big data algorithms are notoriously opaque and incomprehensible, sometimes even to those who are applying them. Two of the largest providers of predictive algorithms in the criminal justice system are corporations who claim that the inner workings of their software are trade secrets.[66]

---

[65] There is a risk to the trustworthiness factor if the human decision-maker claims that she is listening to the defendant's side of the story but in fact is merely blindly following the recommendation of the algorithm.

[66] The companies that make two of the leading predictive algorithms, COMPAS and PredPol, still refuse to make their algorithms public. Since this is valuable proprietary information, police departments foster this atmosphere of secrecy by signing non-disclosure agreements with technology vendors, fighting public record requests, and cooperating with federal officials to protect information from disclosure. *See* RIEKE ET AL., *supra* note 36, at 20. The Arnold Foundation, which creates the PSA tool that is used in thousands of pretrial assessments, has released the details of its algorithm. Tashea, *Calculating Crime*, *supra* note 11, at 58. A representative of the

One way to ensure greater transparency would be through legal compulsion. In the past, the Supreme Court has toyed with the idea of requiring transparency in sentencing. In the 1977 case of *Gardner v. Florida*, the defendant was sentenced to death after the judge relied upon a sentencing report that included a confidential section which was not disclosed to the defense attorney.[67] The Supreme Court overturned the conviction, holding that the defendant was denied due process because the death penalty was imposed in part "on the basis of information which he had no opportunity to deny or explain."[68] The Court also noted that "even if it were permissible to withhold a portion of the report from the defendant . . . it would nevertheless be necessary to make the full report a part of the record to be reviewed on appeal." Some defense attorneys were hopeful that *Gardner* represented a new era of transparency in sentencing decisions, but the Court later made clear that its holding was limited to capital cases.[69]

The movement for more transparency gained some momentum after *United States v. Booker* returned massive amounts of sentencing discretion to trial judges in 2005.[70] In the wake of *Booker*, appellate courts were faced with the task of deciding whether each of the sentences handed down by the trial courts were "reasonable" under the Sentencing Reform Act.[71] The Seventh Circuit responded in *United States v. Cunningham* by creating a two-part test to evaluate reasonableness.[72] First, an appellate court had to ensure the sentence was substantively reasonable by determining whether the sentence was consistent with the factors listed in the sentencing statute. Next the appellate court had to conduct a review of the procedural reasonableness of the trial court's decision. Part of this reasonableness review was to determine whether the judge gave appropriate consideration to the defendant's arguments for downward departures, and according to *Cunningham*, this included a duty to explain her

Arnold Foundation argued that "it's important from a fairness perspective for all the parties to understand what goes into a risk assessment." *Id.*

67 Gardner v. Florida, 430 U.S. 349, 353 (1977).

68 *Id.* at 362.

69 *See* O'Dell v. Netherland, 521 U.S. 151, 162 (1997).

70 United States v. Booker, 543 U.S. 220, 245 (2005) (holding that the United States Sentencing Guidelines practice of enhanced sentencing based on the judge's determination of a fact violates the Sixth Amendment right to a jury trial and thus returning significant sentencing discretion to the trial judge).

71 *Id.* at 261-62; *see also* 18 U.S.C. § 3553(a) (2018).

72 United States v. Cunningham, 429 F.3d 673, 675 (7th Cir. 2005).

reasoning on the record.[73] Thus, the Seventh Circuit created a "duty to explain" for sentencing judges.

Unfortunately, this legally mandated duty was short-lived. It was adopted in principle by a few other circuits,[74] but was ultimately rejected by the Supreme Court, which held in a later case that a trial court has no legal duty to expressly address the defendant's arguments for a variance.[75] As one commentator notes, the Court "set the bar [for explaining the sentence] so low that the requirement can seemingly be satisfied in nearly all cases by a bare acknowledgement of the defendant's arguments for a variance."[76]

Cunningham's duty to explain did not rest on the notion of procedural justice or even due process concerns; rather, the purpose of requiring an explanation by the trial judge was to ensure the appellate court had a sufficient record with which to evaluate the sentencing decision. Thus, the justification for *Cunningham* was based on concerns for judicial efficiency in order to ensure that an appellate court could swiftly and competently evaluate a trial court's sentencing decision.[77]

However, the use of big data algorithms at sentencing raises a new issue regarding a "duty to explain" sentencing decisions. Even when judges make decisions without the aid of predictive algorithms, there is a procedural justice benefit to having the trial judge fully justify her sentencing decision.[78] This benefit is increased dramatically if the

---

[73] *Id.* at 679. *Cunningham* did not require trial judges to explicitly discuss every one of the defendant's arguments on the record; frivolous arguments could be silently ignored. The court set out three factors to determine whether an argument needed to be addressed by the trial judge: (1) whether the argument in question was one of the principal ones advanced by the defendant; (2) whether the argument was based on a factor previously recognized in the law as an acceptable basis for a below-guidelines sentence; and (3) whether the reasons for rejecting the argument would have been self-evident to someone who was familiar with the facts of the case. *Id.*

[74] *See, e.g.*, United States v. Cooper, 437 F.3d 324, 329-30 (3d Cir. 2006).

[75] Rita v. United States, 551 U.S. 338, 358-59 (2007). The Supreme Court mentioned that it is always beneficial for the trial court to explain its reasoning at sentencing, but that a "brief" statement of reasons which did not respond to the defendant's primary arguments (as in the facts of *Rita*) was legally sufficient. *Id.* at 357-58.

[76] O'Hear, *Explaining Sentences*, *supra* note 46, at 469.

[77] Ironically, concerns for judicial efficiency were probably the reason that the Supreme Court rejected the duty to explain. *See Rita*, 552 U.S. at 358-59.

[78] *See* Adam Lamparello, *Social Psychology, Legitimacy, and the Ethical Foundations of Judgment: Importing the Procedural Justice Model to Federal Sentencing Jurisprudence*, 38 COLUM. HUM. RTS. L. REV. 115, 157-58 (2006) (arguing on procedural justice grounds for no limits on either side for speaking during the sentencing hearing, and for judges to clearly explain their sentencing decisions).

judge's reasoning cannot be derived implicitly from the facts on the record, but rather is the result of an algorithm that takes dozens or perhaps hundreds of seemingly irrelevant factors into consideration. Thus, the widespread adoption of predictive algorithms in the sentencing process has increased the need for a legal duty to explain, which could be grounded in the same legal reasoning that supported the original *Cunningham* decision.[79]

If predictive algorithms are made more transparent, and if the authority figures who use them take the time to explain them to defendants, then predictive algorithms could represent a substantial increase in procedural justice. They have the potential to be seen as far more neutral and trustworthy than human decision-makers, who are intrinsically biased and who may not always be able to fully explain the reasoning behind their decisions.

III. PUBLIC PERCEPTION OF PREDICTIVE ALGORITHMS

The preceding section discussed how individuals involved in the criminal justice system, specifically defendants, might react to being subject to predictive algorithms. Supporters of predictive algorithms should also be concerned with how the general population reacts to the adoption of predictive algorithms in the criminal justice system. If the general public is uncomfortable with police officers and judges relying on predictive algorithms to make decisions, predictive algorithms will never become widespread regardless of how efficient or fair they may be.

A. Distrusting the Machines

1. Algorithmic Aversion

Certain psychological factors influence how the general public views the use of predictive algorithms. The first is a well-known psychological phenomenon known as algorithmic aversion. Put simply, algorithmic aversion means that people will generally choose predictions made by human beings over those made by computers,

[79] *See* O'Hear, *Explaining Sentences*, *supra* note 46, at 475-76. While Professor O'Hear wrote his article before the widespread adoption of predictive algorithms in the sentencing process, he uses procedural justice arguments to argue for a duty to explain even when the judge is making the decision on his or her own. As he notes, "[a] rational, responsive decisionmaking process conveys a message of respect for the basic human dignity of the individuals affected by the decision, regardless of the ultimate content of the decision." *Id.*

even after they have seen evidence that the computer predictions are more accurate.[80]

Algorithmic aversion occurs when an individual observes an algorithm make a mistake, as will happen with any algorithm, no matter how close to perfect it may be. Studies show that individuals overreact to mistakes made by algorithms, and thus their level of trust in the algorithm decreases far more precipitously than it would if a human decision-maker made a similar error.[81]

For example, assume you are leaving your office and your co-worker warns you of a serious accident on your route home and suggests an alternate route.[82] You take the alternate route, but it takes you twenty minutes out of your way and is clogged by traffic. You later learn that your normal route home had no traffic problems. You may be somewhat less likely to take your co-worker's navigation advice again, but her mistake is unlikely to lead you to conclude that she is untrustworthy. You realize that people sometimes make mistakes, and in time you are likely to forget about the incident altogether.

Now instead assume that you never talked to a co-worker, but rather decided to rely on a new navigation app that you downloaded that day. The app tells you that there is a serious accident on your normal route home and directs you along an alternate route. As in the previous hypothetical, the alternate route adds twenty minutes to your drive and is clogged with traffic, and you discover later that your normal route did not in fact have a traffic jam. You are much more likely to lose faith in the new navigation app; in fact, you may never use it again. In other words, "the errors that we tolerate in humans become less tolerable when machines make them."[83]

This psychological tendency was confirmed in a recent experiment in which researchers provided subjects with information about applicants for business school and then asked them to predict how well each of the applicants would perform in a MBA program.[84] For the first set of data, the subjects were able to make their own

---

[80] *See* Dietvorst, Simmons & Massey, *supra* note 16, at 1. The authors point to a number of potential reasons for algorithmic aversion, including "the desire for perfect forecasts, the inability of algorithms to learn, the presumed ability of humans to improve through experience, the notion that algorithms are dehumanizing, the notions that algorithms cannot properly consider individual targets, concerns about the ethicality of relying on algorithms to make important decisions, and the presumed inability of algorithms to incorporate qualitative data." *Id.* at 2 (citations omitted).

[81] *Id.* at 10-11.

[82] This example is adopted from the Dietvorst, Simmons & Massey paper. *Id.* at 2.

[83] *Id.* at 2.

[84] *Id.*

predictions, and some of them also saw the predictions made by a computer algorithm that was using the same data.[85] On average, the predictive algorithm outperformed the human forecasters by a significant margin.[86] For the next set of data, researchers offered the subjects a monetary bonus for each correct prediction, and allowed them to either make their own predictions or choose the predictions made by the computer algorithm.

The experiment was run five different times. In each study, the subjects were more likely to choose their own predictions than those of the computer algorithm when their bonuses were tied to the accuracy of their predictions.[87] This was even true when the subjects had seen the results of their own predictions and the results of the computer algorithm from the first set of data. In fact, subjects were *more likely* to choose their own predictions over those of the computer algorithm after seeing the results of both methods, even when the computer algorithm demonstrated twice as much accuracy as their own predictions.[88] The findings indicated that when the subjects got to choose between another human's predictions and the predictions of the computer algorithm, the subjects would choose the other human's predictions even when they saw that the computer algorithm's predictions were more accurate.

The authors of the study concluded that the subjects who saw the algorithm perform were more influenced by the mistakes the algorithm made than by its correct predictions. This was true even though the subjects also saw the human predictions err, and even though they saw the humans err more often. The authors conclude that "people are quicker to abandon algorithms that make mistakes than to abandon humans that make mistakes, even though, as is often the case, the humans' mistakes were larger."[89] Thus, the more a person

---

[85] *Id.* at 3-4. The subjects were told that the predictive algorithm was a statistical model that used the same data and was designed "by thoughtful analysts." *Id.* at 4.

[86] For predicting which applicants would do best at business school, humans produced fifteen to twenty-nine percent more error than the predictive algorithm, and for the other task (predicting the number of airline passengers that depart from each state), humans produced ninety to ninety-seven percent more error. *Id.* at 5-6.

[87] *Id.* at 6.

[88] *Id.*

[89] *Id.* The authors of the study hypothesized that:

> [S]ome magnitude of advantage must lead participants who see the algorithm perform to be more likely to choose it — for example, if they were to see the algorithm predict all outcomes exactly right — the model's large advantage in these studies was not large enough to get them to do so.

is exposed to the decisions of predictive algorithms, the less likely the person is to trust those decisions.

These findings imply that the more widespread predictive algorithms become in the criminal justice system, the more resistance they will encounter from the general public, since the general public will inevitably see the failures of predictive algorithms along with their successes. This phenomenon will only be accelerated by the way that these predictive algorithms are covered in the media. Currently, media stories on predictive algorithms in the criminal justice system fall into one of three categories. First are stories that criticize the use of predictive algorithms based on the algorithm's racial bias[90] or lack of transparency.[91] These stories generally argue against the widespread adoption of predictive algorithms unless they are reformed. A second category of stories are those written by quantitative reporters, such as those found at the New York Times' Upshot[92] or fivethirtyeight.com.[93] Stories in this category are more nuanced, often reporting on studies that demonstrate the accuracy of these tools but acknowledge their limitations and potential biases. Some also note that much of the bias seen in the algorithms' results is reflective of pre-existing bias in the criminal justice system.[94]

The third category of media coverage on criminal justice algorithms focuses on individual cases in which a particular algorithm has failed. These stories use individual cases to attack predictive algorithms from both sides of the political spectrum. They could highlight an

Id.

[90] *See, e.g.*, Angwin, *supra* note 22 (arguing that the COMPAS system is "remarkably unreliable" and results in "significant racial disparities").

[91] *See* Adam Liptak, *Sent to Prison by a Software Program's Secret Algorithms*, N.Y. TIMES (May 1, 2017), https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html ("There are good reasons to use data to ensure uniformity in sentencing. It is less clear that uniformity must come at the price of secrecy, particularly when the justification for secrecy is the protection of a private company's profits.").

[92] *See* Sam Corbett-Davies, Sharad Goel & Sandra Gonzalez-Bailon, *Even Imperfect Algorithms Can Improve the Criminal Justice System*, N.Y. TIMES (Dec. 20, 2017), https://www.nytimes.com/2017/12/20/upshot/algorithms-bail-criminal-justice-system.html.

[93] *See* Anna Maria Barry-Jester, Ben Casselman & Dana Goldstein, *Should Prison Sentences Be Based on Crimes That Haven't Been Committed Yet?*, FIVETHIRTYEIGHT (Aug. 4, 2015), https://fivethirtyeight.com/features/prison-reform-risk-assessment/.

[94] *See, e.g.*, Corbett-Davies, Goel & Gonzalez-Bailon, *supra* note 92 ("It is not biased algorithms but broader societal inequalities that drive the troubling racial differences we see in Broward County and throughout the country. It is misleading and counterproductive to blame the algorithm for uncovering real statistical patterns. Ignoring these patterns would not resolve the underlying disparities.").

individual case in which a defendant was classified as "high risk" at a bail or sentencing hearing but then never committed another crime,[95] or detail a case in which a defendant is released based on an algorithm's recommendation and then commits another violent crime.[96]

This third category of stories exacerbates the effects of algorithmic aversion. To return to the self-driving car analogy, stories of self-driving cars running over pedestrians will always have a greater impact on public opinion than stories that report statistics about the overall safety of self-driving cars. Likewise, stories about defendants who are released on bail based on the recommendation of a predictive algorithm and who then commit murder the following week will disproportionately influence public opinion against using these algorithms.

2. The Need for Human Input

A second reason that the general population distrusts computer-driven predictive algorithms is the belief that "human input" is necessary in significant decisions. A recent Pew Research Center survey of over 1,300 technology experts, scholars, corporate practitioners and government leaders found that one of the primary concerns with the spread of algorithms in any field is that they tend to dehumanize the decision-making process.[97] Decisions are no longer considered to be "real, thinking, feeling, changing beings;" as they are merely a piece of data for an algorithm to manipulate.[98] As one respondent reported: "This has been going on since the beginning of the industrial revolution. Every time you design a human system optimized for efficiency or profitability you dehumanize the

[95] *See* Angwin, *supra* note 22.

[96] *See* Eric Westervelt, *Did a Bail Reform Algorithm Contribute to this San Francisco Man's Murder?*, NPR: ALL THINGS CONSIDERED (Aug. 18, 2017), https://www.npr.org/2017/08/18/543976003/did-a-bail-reform-algorithm-contribute-to-this-san-francisco-man-s-murder. The story details the murder of Edward French, who was allegedly killed by two individuals, one of whom, Lamonte Mims, had been arrested mere days before for gun possession and parole violation. Mims was given a low-risk score by the PSA algorithm designed by the Arnold Foundation and released without bail. Apparently, the pretrial division inputted incorrect information into Mims' PSA data, resulting in an incorrect score, but critics of the algorithm argued that the judge gave the PSA assessment too much weight without exercising her own judgment. *Id.*

[97] Rainie & Anderson, *supra* note 3.

[98] *Id.*

workforce . . . . When you remove the humanity from a system where people are included, they become victims."[99]

The concern about dehumanization is even more pronounced in the criminal justice field. Unlike other areas of law, criminal law carries with it a moral judgment that the perpetrator has broken the social contract,[100] and it is inappropriate for machines to make that moral judgment. Yet predictive algorithms are now determining where police officers patrol, whom they arrest, whether a defendant remains free before trial, and how long a defendant should be sentenced after he is convicted. Allowing machines to make these fundamental decisions about human liberty can be perceived as an affront to human dignity.[101] Predictive algorithms cannot show mercy; they are like an automated vacuum cleaner that sucks up the dust and the crickets without caring or even realizing the difference.[102] When a police officer or a judge makes a mistake, observers can choose to forgive, blame, or feel some similar cathartic emotion towards the human being who made the decision. When a predictive algorithm makes a mistake, there is no easy way to deal with the injustice that was created.

Professor Andrea Roth raised this concern when discussing "mechanized criminal adjudication."[103] She argues that the criminal justice system needs individualized human judgment to "fully assess blameworthiness through a combination of complex fact-finding, equitable discretion, and mercy."[104] Although machines may be better at complex fact-finding, they generally cannot be programmed to consider questions of equitable discretion and mercy. Professor Stephen Henderson echoed this concern in a recent article in which he proposed rules for adopting big data in the criminal justice system.[105] One of his rules was that the ultimate decider must always be

---

[99] *Id.*

[100] *See* Henry M. Hart, Jr., *The Aims of the Criminal Law*, 23 LAW & CONTEMP. PROBS. 401, 404-05 (1958).

[101] *See* Rainie & Anderson, *supra* note 3 (surveying individuals regarding their concerns with the increasing use of algorithms where one respondent stated, "[i]ndividual human beings will be herded around like cattle, with predictably destructive results on rule of law, social justice and economics").

[102] Thanks to Professor Elizabeth Joh for this analogy.

[103] Andrea Roth, *Trial by Machine*, 104 GEO. L.J. 1245, 1247 (2016).

[104] *Id.* Professor Roth argues that "[c]alls for robot judges and juries are typically met with derision" because machines are unable to take into account "softer" goals of the criminal justice system such as dignity, equity, and mercy. *Id.*

[105] Stephen E. Henderson, *A Few Criminal Justice Big Data Rules*, 15 OHIO ST. J. CRIM. L. 527, 530 (2018).

human.[106] He argues this is the case even if, as he concedes, human-influenced decisions are often less correct than fully automated decisions. Rather, he proposes this rule because "on some intuitive level, it seems important to our humanity."[107]

The most extreme example of this potential dehumanization in the criminal context would be to replace a jury with a computer that could receive all the evidence and then determine whether to convict or acquit the defendant. Such a scenario seems unfathomable, not because we cannot imagine the technology, but because we cannot accept a computer making a final decision regarding guilt or innocence. But using a predictive algorithm to calculate a prison sentence is not so different from using an algorithm to determine guilt. In fact, there is a greater need for considerations of equity, mercy, and human dignity at sentencing than at trial. Deciding guilt or innocence is a purely factual matter,[108] while determining a sentence requires the sentencing authority to exercise discretion.

To a lesser degree, predictive algorithms take away discretion from police officers, as well. Individual police officers wield tremendous discretion as to when to make an arrest. Often, they may see a minor crime occurring (such as trespassing in a park after hours or smoking a marijuana cigarette) and issue a warning rather than make an arrest. The more their duties are determined by computers, the fewer opportunities they will have to exercise this discretion.

The obvious solution to this problem is to follow Professor Henderson's suggestion and ensure that a human being is always the final decider, with the authority to follow, ignore, or amend the recommendation of the predictive algorithm. The paradox is that keeping a human being in the loop removes one of the primary benefits of using predictive algorithms: their complete disregard of irrelevant subjective factors. A judge who shows mercy to one defendant but not another is, in a very real sense, being unfair to other defendants who she deemed unworthy of mercy. The judge's own preconceptions of who is worthy of a second chance, based perhaps on how the defendant conducts himself in court or the convincing testimonials from a family member, are much more likely to be inaccurate than the empirically tested factors used by a computer. (They may also be based, consciously or unconsciously, on race or religion or economic class.) And the reasons that a police officer might

[106] *Id.* at 533.

[107] *Id.* at 534.

[108] Assuming one discounts the rare and controversial practice of jury nullification.

look the other way when discovering criminal activity are probably not linked to any evidence as to who is likely to re-offend. In short, as soon as you address the intuition that many people have about needing human input in the system, you invite back all of the inaccuracies and biases that predictive algorithms are meant to eliminate.

3. Using Broad Group Characteristics to Determine the Fate of an Individual

A third reason that the general population may feel uncomfortable with the use of predictive algorithms is that algorithms inevitably use statistics about groups as a factor in determining what should happen to an individual. Former United States Attorney General Eric Holder raised this concern in an interview in 2014, in which he stated that "[u]sing group data to make an individualized determination . . . can result in fundamental unfairness."[109]

One article gives the example of Milton Fosque, who was convicted of his third drunk driving offense in Pennsylvania. Pennsylvania had just adopted a risk assessment program for sentencing, and if Fosque had been sentenced under the new program, his specific characteristics would have been translated into points, which would then determine his sentence For example, the fact that he is male and lived in an urban area would increase his risk factor by two points, while his four prior arrests would increase his risk factor by four more points.[110] As a result of these elevated risk factors, Pennsylvania's predictive algorithm would have calculated that he had a forty-nine percent chance of re-offending.[111] When told about this system, Fosque objected to this calculation, claiming that the predictive algorithm would not consider enough of his own individual characteristics, such as how long he had remained sober since his arrest.[112] His reaction to the risk assessment tool sums up what many

---

[109] Joshua Barajas, *Holder: Big Data Is Leading to 'Fundamental Unfairness' in Drug Sentencing*, PBS NEWS HOUR (July 31, 2014, 6:29 PM), https://www.pbs.org/newshour/politics/holder-big-data-leading-fundamental-unfairness-drug-sentencing; *see also* Starr, *Sentencing, by the Numbers*, *supra* note 33 (discussing a letter Attorney General Holder also sent in 2014 to the United States Sentencing Commission criticizing this practice).

[110] Anna Maria Barry-Jester, Ben Casselman & Dana Goldstein, *The New Science of Sentencing*, MARSHALL PROJECT (Aug. 4, 2015, 7:15 AM), https://www.themarshallproject.org/2015/08/04/the-new-science-of-sentencing.

[111] *Id.*

[112] *See id.* Fosque argued that because of his newfound sobriety, the chance of him

Americans might say: "You mean to tell me they're using statistics to determine what's going to happen to me? That ain't right."[113]

The use of group statistics to predict the future behavior of an individual raises two separate issues. The first is whether it is unfair to use group statistics at all, rather than focusing on the individual himself. This seems to be what Attorney General Holder and defendants like Milton Fosque find concerning. However, courts have always considered these broad factors in making their bail and sentencing determinations. The concern is that predictive algorithms give these factors a "scientific" veneer, so that judges give them too much weight and rely less on individual characteristics that are not taken into consideration by the algorithm, such as the fact that an individual convicted of drunk driving has been sober for a year and goes regularly to Alcoholics Anonymous meetings, or that a serial burglar has recently converted to Christianity.[114] Of course, these individualized factors could also lead a judge to *increase* bail or a sentence. A defendant may act in a way that a judge finds irrational or troubling during the hearing or a defendant may exhibit signs of mental instability.[115] Thus, if a judge is discouraged from using individualized factors because of the recommendation of a predictive algorithm, the effect on a defendant could be positive or negative.

Once again, supporters of predictive algorithms argue that suppressing a judge's consideration of individualized factors is a feature of predictive algorithms, not a bug. One of the purposes of using risk assessment tools is to remove subjective factors from a judge's decision-making process, since those factors are often based on unconscious biases and are less likely to be accurate than the factors used by the algorithm. The fundamental nature of these tools is that they use statistics based on group characteristics to predict the future behavior of individuals, and that doing so actually increases both the accuracy and the fairness of the process.

---

re-offending was zero. *Id.*

[113] *Id.* The judge sentenced Fosque without using the algorithm and gave Fosque one year in jail (out of a possible one to five) and five years of probation. *Id.*

[114] This is a concern even though the judge may retain the discretion to deviate from the algorithm's suggestions. Although the judge will have the ability to consider factors extraneous to the algorithm, she may be reluctant to do so if she knows that the algorithm has a high degree of accuracy.

[115] *See, e.g.*, *Serial: You've Got Some Gauls*, CHI. PUB. RADIO (Sept. 20, 2018), https://serialpodcast.org/season-three/2/youve-got-some-gauls (documenting a judge in Cleveland, Ohio, who sentences, at least in part, based upon his perceptions of the individual, whether positive or negative).

A more specific problem is that some of these tools use group statistics based on the individual's *status* in addition to the individual's conduct. Professor Sonja B. Starr, an outspoken critic of using risk assessments during sentencing, makes this distinction in her critiques. She agrees that an individual's past criminal conduct is a legitimate factor;[116] for example, if the statistics show that an individual with three prior convictions for burglary is twice as likely to commit another burglary than an individual with a clean record, it would be fair to use a defendant's prior convictions as a factor in his risk assessment. But Professor Starr notes that many of the factors that increase a defendant's risk score, such as gender, age, low levels of education, financial history, and family members' criminal history, are not based on the defendant's conduct but based on his status.[117] She argues that relying on these factors "contravenes the principle that punishment should depend on what a defendant did, not on who he is or how much money he has."[118]

As noted in Part I, the problem of over-reliance on group characteristics is not just a problem of perception; the requirement of individualized suspicion is necessary to ensure that the defendant's constitutional rights are being preserved.[119] As long as predictive algorithms consider some individualized factors in making their determinations, they can alleviate this problem to some extent.

B. Trusting the Machines

Although human beings demonstrate certain irrational psychological barriers to accepting predictive algorithms, they also frequently exhibit the opposite characteristic: the tendency to overrate the validity of machine decisions, known as "automation bias."[120] Studies have shown that machine results are often seen as "error-resistant,"[121] and that a recommendation by a computer program will

116 Starr, *Sentencing, by the Numbers*, *supra* note 33.

117 *Id.* Not incidentally, many of these high-risk status designations are disproportionately prevalent among non-whites and individuals with low socio-economic status. *See id.*

118 *Id.*

119 *See supra* notes 31–35 and accompanying text.

120 Linda J. Skitka et al., *Automation Bias and Errors: Are Crews Better than Individuals?*, 10 INT'L J. AVIATION PSYCHOL. 85, 85-86 (2000).

121 *See* Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1254 (2008); Raja Parasuraman, *Designing Automation for Human Use: Empirical Studies and Quantitative Models*, 43 ERGONOMICS 931, 936 (2000); *see also* Thomas B. Sheridan, *Speculations on Future Relations Between Humans and Automation*, *in*

often be seen as infallible.[122] In her recent article, Professor Roth traces the history of this pro-technology bias in the criminal justice system, from intoxilyzers to DNA analysis, and notes that even the most advanced technology is based on human programming and input.[123] In other words, just as algorithmic avoidance leads people to irrationally reject machine decisions because they will put too much weight on high profile mistakes, automation bias will lead people to irrationally accept machine decisions because they put too little weight on the fallible human inputs into the system.

As noted in the introduction, people *do* accept and even embrace algorithms in countless aspects of their daily life. Algorithmic aversion aside, people willingly choose to follow the suggestions and recommendations of computers in a wide variety of fields. We use navigation apps on our phone and shop for airline tickets using automated fare finders, even though paper maps and human travel agents are still available. Most of us are at least dimly aware that algorithms power much of our modern world, from routing our cell phone calls and our plane fights to directing our internet searches and our social media feeds, and we still eagerly use these algorithms because of their convenience and accuracy.

But as we saw in the previous section, using algorithms in the criminal justice context is different. Concerns for human dignity and a desire for a more individualized process make people less likely to accept machines making decisions that affect human liberty. Thus, the general public will be more willing to agree with predictive algorithms in some criminal law contexts than in others. There are many different aspects of the criminal justice system that are using or could potentially use predictive algorithms.[124] Perhaps it is more acceptable to have computer programs tell police where to patrol than to tell judges how long a defendant's sentence should be.[125] One area that

AUTOMATION AND HUMAN PERFORMANCE: THEORY AND APPLICATIONS 449, 458 (Raja Parasuraman & Mustapha Mouloua eds., 1996) ("It is so tempting to trust to the magic of computers and automation. . . . [I]f a computer program compiles, we often believe, the software is valid and the intention will be achieved.").

122 M.L. Cummings, *The Social and Ethical Impact of Decision Support Interface Design*, *in* INTERNATIONAL ENCYCLOPEDIA OF ERGONOMICS AND HUMAN FACTORS 1249, 1250 (2d ed. 2006).

123 Roth, *supra* note 103, at 1269-76.

124 *See id.* at 1253-69 (providing an overview of how some machines provide evidence, decide guilt, and determine punishment).

125 It would also be interesting to determine whether there was a positive (or negative) correlation between the contexts in which predictive algorithms are *perceived* to be more legitimate and the areas in which they are actually more fair.

deserves further study is *why* people feel that humans should remain in the loop, so that we can determine what type of decisions can be automated without creating controversy.

The least objectionable use of computer algorithms is when machines are not making decisions on behalf of human beings, but merely creating a fairer context within which the humans will make their decisions. Professor Andrew Ferguson has proposed using predictive algorithms to help construct a jury pool that perfectly matches the demographics of a community.[126] Under this regime, big data would supplement the voir dire questions and peremptory challenges that are now used to ensure a fair jury.

The next step would be to allow computer algorithms to provide information to humans in order to aid them in making their decisions. Professor Andrea Roth offers a number of suggestions for using computer algorithms during a trial, such as text analysis software to determine whether a witness is lying on the stand, or a Watson-like computer that can fact-check witness' assertions in real time.[127] But this becomes a slippery slope. When a predictive algorithm tells the police which neighborhoods are likely to have criminal activity in the next hour, or when it tells a judge whether or not a defendant is a "high risk" for flight, it is merely providing information so that the human can make a better decision, but if the human decision-maker routinely follows the computer recommendation, it could appear as though the human has been taken out of the loop.

To measure how the general population views the use of predictive algorithms in the criminal justice system, I designed a survey to measure how individuals would respond to the use of a predictive algorithm in the context of a bail hearing. I chose the bail hearing because the use of predictive algorithms to assess risk of flight is one of the most common uses of predictive algorithms in the modern criminal justice system. And in order to follow Professor Henderson's advice from the previous section to always keep a human in the loop, the survey keeps a human judge as the ultimate decider, though the frequency with which the judge agrees with the predictive algorithm is kept intentionally vague.

The survey was administered through the Mechanical Turk marketplace, a research tool used with increasing frequency by social science researchers,[128] which provides respondents who are more

[126] *See* Andrew Guthrie Ferguson, *The Big Data Jury*, 91 NOTRE DAME L. REV. 935, 940-41 (2016).

[127] Roth, *supra* note 103, at 1297.

[128] Christopher T. Robertson et al., *Organs and Inducements: Perceptions of Efficacy,*

representative of the population at large than the usual participants on academic surveys, such as college students.[129] The survey was designed to measure whether using a predictive algorithm to advise a judge regarding a bail decision affects the perceived fairness of the process. Rather than asking individuals about this question directly, the survey provided respondents with a hypothetical defendant and then randomly presented half the respondents with a bail hearing in which a judge set bail based on his or her own judgment, and half the respondents with a bail hearing in which predictive computer software gave the judge a recommendation and the judge accepted the recommendation. Respondents who were presented with a human judge who did not use an algorithm were told that the human judge "had over ten years of experience conducting bail hearings." Respondents who were presented with a human judge who accepted the recommendation of the predictive algorithm were told that the algorithm had "a high degree of accuracy in predicting which defendants are likely to appear in court and which are likely to not appear in court."

The respondents were asked whether they believed the bail determination process was fair to the defendant and whether they agreed with the level of bail that was set. In examining the results, we can compare the responses of the respondents who were presented with a judge advised by the predictive algorithm with the results of the respondents who were presented with a judge who acted without any machine assistance. If both groups demonstrated equal levels of satisfaction with the results of the bail hearing, we can conclude that the use of an algorithm did not decrease the perceived legitimacy of the process. On the other hand, if respondents who saw the predictive algorithm assist the judge demonstrated a higher level of dissatisfaction with the process, we can conclude that the use of the algorithm decreases the perceived legitimacy of the process.

The hypothetical case presented by the survey involved a defendant at a bail hearing who had been arrested for armed robbery. Respondents were given numerous details about the defendant's background and the evidence against him.[130] The respondents were

Morality, and Politics of Potential Cadaveric Organ-Transplantation Reforms, 77 LAW & CONTEMP. PROBS. 101, 108 n.30 (2014).

[129] *See* Danielle N. Shapiro et al., *Using Mechanical Turk to Study Clinical Populations*, 1 CLINICAL PSYCHOL. SCI. 213, 213 (2013). Respondents from the Mechanical Turk can skew younger, more female, and more educated than the general population. Jeffrey Bellin, *The Silence Penalty*, 103 IOWA L. REV. 396, 411 (2018).

[130] During the first court case, no bail was set and he made every court appearance.

divided into four groups of 150 people each.[131] The first group ("No machine/low bail") involved the judge setting a relatively low bail ($4,000) without any input from a computer program; [132] the second ("Machine/low bail") involved the judge following the advice of a computer program and setting a relatively low bail ($4,000); the third ("No machine/high bail") involved the judge setting a relatively high bail ($20,000) without any input from a computer program; and the fourth ("Machine/high bail") involved the judge following the advice of a computer program and setting a relatively high bail ($20,000).

The results are as follows:

	No machine/ low bail ($4k) (n=142)	Machine/ low bail ($4k) (n=147)	No Machine/ high bail ($20k) (n=145)	Machine/ high bail ($20k) (n=144)
Median Average Bail Suggested by Respondents	$80,000	$8,000	$20,000	$20,000

During the second court case, bail was set at $2,000, and he posted bail and then failed to appear. He was later arrested on a bench warrant and then pled guilty. In the current case, the defendant had been arrested for an armed robbery. He had been arrested near the scene of the robbery only a few minutes after the crime; the victims identified him as the perpetrator; and the police recovered a gun and the victim's wallets when they searched him incident to arrest. Respondents were told that the judge is now setting bail for this third crime. The entire text of the survey, along with the four different scenarios that were used and the questions that were asked, is reproduced in Appendix A.

[131] The total number of responses for each question is lower than 150 for each group, since not every respondent answered each question.

[132] Before the main survey was run, a control group was given this same hypothetical case without any mention of a judge setting bail and then asked what bail the members of the control group would set if they were the judge. The median average answer from this group was $10,000, and the mean average was $18,744. Thus, in designing the main survey, we set the "low bail" amount at well below the median and the "high bail" amount at will above the median. This was intended to ensure that almost all the respondents would disagree with the amount of bail that was set, and therefore be a better test of whether they approved of the procedure that was followed.

Mean Average Bail Suggested by Respondents	$14,816	$14,014	$32,793	$34,535
Standard Deviation	20,312	24,215	37,600	35,697
Bail set by court was far too low	17%	14%	10%	8%
Bail set by court was too low	44%	43%	21%	24%
Bail set by court was about right	29%	35%	51%	49%
Bail set by court was too high	7%	7%	15%	13%
Bail set by court was far too high	1%	1%	3%	4%
Defendant was treated very fairly in the process	59%	50%	52%	53%
Defendant was treated somewhat fairly in the process	25%	30%	29%	30%
The process was neither fair nor unfair to the defendant	11%	13%	11%	7%

The defendant was treated unfairly in the process	2%	5%	5%	7%
The defendant was treated very unfairly in the process	1%	1%	1%	1%

We will first consider the bail amounts suggested by the respondents. When respondents were informed that a judge had set bail at a certain amount, either with or without a computer recommendation, that information had an anchoring effect, influencing the amount of bail that respondents thought was appropriate. As we can see from first level of results, this anchoring effect was identical whether or not a predictive algorithm was involved in the process. A low bail decision lowered the median bail amount to $8,000 regardless of whether a computer assisted in the process, and a high bail decision raised the median bail to $20,000 (in this case, identical to the bail set by the judge) in both scenarios. Thus, respondents found the bail set by the judge to be equally influential regardless of whether the judge used a predictive algorithm in making his or her decision.

Likewise, when participants were asked whether they believed the bail set by the judge was appropriate, similar numbers of respondents agreed with the bail regardless of whether a predictive algorithm was advising the judge. And when asked whether the process was fair, eighty-four percent and eighty-one percent of the "no machine" groups replied that it was either "fair" or "very fair," while eighty percent and eighty-three percent of the "machine" groups replied this way.

The conclusions from the survey are limited by a number of factors. First, as with most surveys, the respondents knew they were responding to a hypothetical case; their reactions to an actual criminal case may have been different. Second, the survey intentionally presented a "best case scenario" for convincing people of the legitimacy and fairness of predictive algorithms. The survey explained how the algorithm worked (that it gathered and compared data from "hundreds of thousands of cases"), and that algorithm had a "high degree of accuracy." Unlike the respondents in the algorithmic aversion study, the respondents here did not get to see the algorithm in action on multiple cases, and thus never saw the algorithm fail. Had respondents been exposed to data which showed the predictive

algorithm correctly predicting flight risk four times out of five, but incorrectly predicting it one time out of five, they may have shown less faith in the algorithm's recommendation, even if they had also been told that the computer is more accurate than the human judge. Finally, in the hypothetical case involving the predictive algorithm, the human judge retained the decision-making authority: the text of the survey stated that "the judge accepts the program's assessment," indicating that the judge had the power to reject the assessment. As of now, no jurisdiction allows computers to make the ultimate decision in bail or sentencing hearings, and such a possibility seems unlikely, since the public may never be ready to take humans out of the loop on these decisions. However, it is worth noting that the results of the survey may not hold if there was no human judge involved at all.

Even given these limitations, these results all strongly indicate that the general public is willing to accept predictive algorithms in this context. The survey showed no evidence of a widespread dissatisfaction or anxiety about judges relying on predictive algorithms in this context. The comments made by the respondents support this conclusion. The survey ended with the open-ended question "Why do you think the process was fair or unfair?" A total of sixty respondents mentioned the computer in their response,[133] and forty-three of these sixty (seventy-two percent) indicated that the presence of the predictive algorithm increased their perception of fairness. Some representative comments included:

> "It was fair because it [was] evaluated by a machine rather than a person."
>
> "Because the judge used a computer program and there couldn't have been bias."
>
> "It seems fair because the computer has no prejudices and computes the estimate based on facts."

Only fourteen of the responses (twenty-three percent)[134] mentioned the computer in a negative way. Representative comments from this group included:

---

[133] Of the 300 respondents who were given the "machine" scenario, 286 responded to this open-ended question. Sixty of these 286 responses mentioned the use of the computer algorithm. Of course, none of the 300 respondents who were given the "no machine" scenario mentioned the predictive algorithm, since the algorithm did not exist in their hypothetical.

[134] Three of the comments that discussed the computer program (the remaining five percent) were neither positive nor negative about its use.

> "I think the judge has to think for himself instead of using a program with no human thought connected to it."
>
> "Computer generated may not take into account extenuating circumstances."
>
> "I think the bail amount was more than fair for armed robbery but I'm still a little hesitant to be letting a computer determine things like that."

Thus, of those who commented on the use of the computer, three times as many respondents thought the influence of the computer was a positive factor as thought the opposite.

What is perhaps the most telling of all, however, is that the vast majority of individuals who read one of the scenarios in which the judge relied heavily on a predictive algorithm did not even mention the algorithm at all in their comments — there were 286 total comments from the groups that were exposed to the predictive algorithm, and 226 of them (seventy-nine percent) did not think that the use of the predictive algorithm was even worth mentioning. This is consistent with our conclusion that the general population is comfortable with the concept of computers assisting in these decisions, at least as long as a human judge makes the ultimate decision.

These conclusions are mostly consistent with the only other study done so far about the perceived legitimacy of predictive algorithms in the criminal justice system. A.J. Wang of Yale Law School ran a study of 3,369 respondents and asked them to rank the fairness of judicial decision-making when the judge consults a predictive algorithm, sentencing guidelines, or a psychologist.[135] Wang notes that when given no information about the accuracy of the predictive algorithm, respondents rank using algorithms as significantly less fair than using guidelines or consulting psychologists.[136] However, he also found that respondents generally believed that algorithms are significantly less

[135] A.J. Wang, Procedural Justice and Risk-Assessment Algorithms 5-6 (June 21, 2018) (unpublished manuscript) (on file with author). In the paper, Wang argues that his study presents "evidence to the contrary" of the results in my own study. *Id.* at 4. Wang's study is in some ways more comprehensive than my own, with many more respondents and a number of different experiments. For most of the experiments, he finds a strong disapproval of predictive algorithms; however, as noted below, he finds a preference for algorithms when they are described as more accurate than other methods. *Id.* at 7.

[136] *Id.* at 7-10.

accurate than the use of guidelines or psychologists.[137] When his respondents were told that algorithms are more accurate than guidelines or psychologists, they preferred algorithms sixty-six percent of the time.[138]

This is consistent with the results of my own survey, in which all respondents were told that the algorithm has "a high degree of accuracy in predicting which defendants are likely to appear in court and which are likely to not appear in court." It also implies that if respondents are *not* given this information about the high accuracy of algorithms, they are likely to find algorithms to be less fair than other options. Thus, the future legitimacy of predictive algorithms in the criminal justice system may depend on whether the general population believes in the increased accuracy that they can deliver.

CONCLUSION

The integration of predictive algorithms into the criminal justice system has the potential to create a revolutionary change: they can improve the quality of judgments made by police officers and judges; save scarce law enforcement and judicial resources; and hopefully even reduce or remove the irrationality and bias that plague the system. But they also represent a paradigmatic shift in the way decisions are made in the criminal justice system, and as such their adoption could delegitimize criminal justice institutions, both in the eyes of criminal defendants and of the general population.

This Article has shown that human beings have a deep ambivalence about the use of predictive algorithms in general. We accept them in many aspects of our life, and the results of our survey show that most people are perfectly comfortable with predictive algorithms in some areas of the criminal justice system. But even if we provide more

[137] *Id.* at 11.

[138] *Id.* at 11-12. Notably, Wang also found that respondents disapprove of algorithms when the algorithms are equally accurate as other methods. *Id.* at 13. As he points out in his conclusion:

> If [respondents] are told that algorithms are more accurate than other procedures, they balance their desire for the most accurate procedure against their dislike for algorithms. The result is a preference for algorithms when they are the more accurate procedure and only when they are the most accurate procedure. To the extent that algorithms are or become more accurate than existing procedures, policymakers should address and refute public beliefs. If this is done, algorithms become more acceptable.

Id. at 20.

information about the accuracy and efficiency of algorithms, the experiments regarding algorithmic aversion indicate that only a few isolated stories of mistakes by a predictive algorithm can lead people to reject their use.

This Article also suggests a number of steps can be taken to improve the perceived legitimacy of these tools. Individuals must be made aware of the high level of accuracy that predictive algorithms can provide. And predictive algorithms must be transparent in the data they use and the methods they employ. Transparency will enhance the procedural justice experienced by defendants who are subject to their decisions, since it will increase the algorithm's trustworthiness and neutrality. Those who design and use predictive algorithms also need to ensure that the source data used by these tools is untainted by past discriminatory practices, and that each recommendation made by predictive algorithms is based at least in part on the defendant's individualized conduct. Finally, no matter how accurate and fair predictive algorithms are, we will always need to keep a human in the loop as the ultimate decision-maker.

APPENDIX A: TEXT OF SURVEY

John Smith is 25 years old, he is a high-school dropout, and he is employed as a construction worker. He lives with his girlfriend, whom he has been dating for 8 months. He drinks regularly but otherwise has no substance abuse problems. He has lived in the same community for his entire life.

Smith has two prior criminal convictions:

When he was 19 years old, he stole a laptop computer from an electronics store. In that case, no bail was set and he made three court appearances before pleading guilty to a misdemeanor and receiving a sentence of three months in jail.

When Smith was 22 years old, he stole a car from a parking lot. In that case, the judge set bail at $2,000. He paid the bail, and he made two court appearances before failing to appear on the date of his trial. An arrest warrant was issued, and he was arrested and held in jail for two more weeks before he pled guilty to grand larceny. He received a sentence of two years in prison.

John Smith is now charged with armed robbery and unlawful possession of a firearm. According to the police report, he attacked a young couple at night walking home from dinner, pointed a gun at them, and demanded their wallets. The couple gave him their wallets, which included all of their credit cards and a total of about $250 in cash. The couple also told the police that they were in fear for their life when Smith pointed the gun at them.

The couple reported the crime to the police immediately, and fifteen minutes later Smith was arrested a few blocks away. The police found both of the victims' wallets in his pocket, as well as an unloaded gun. Smith did not have a permit to carry a concealed weapon. Smith faces a sentence of between two to five years if he is convicted of this crime.

[Respondents then randomly saw one of the following four scenarios:]

A. Smith is now in front of a judge at a hearing to determine whether bail should be set for this most recent arrest. According to state law, bail should be set at the lowest possible level that is necessary to ensure that the defendant returns to court for trial. There are no bail bondsmen in this jurisdiction; thus, if bail is set, Smith or his family would have to pay the entire amount of whatever bail is set. The judge has over 10 years of experience conducting bail hearings. In determining bail in Smith's case, the judge considers all of the relevant factors, including the severity of the crime, the strength of the evidence, the defendant's prior criminal record, and his background (such as his employment history, family history, and other personal

information). The judge ultimately determines that Smith is not a high risk for flight. The judge sets bail at $4,000.

B. Smith is now in front of a judge at a hearing to determine whether bail should be set for this most recent arrest. According to state law, bail should be set at the lowest possible level that is necessary to ensure that the defendant returns to court for trial. There are no bail bondsmen in this jurisdiction; thus, if bail is set, Smith or his family would have to pay the entire amount of whatever bail is set. The judge has over 10 years of experience conducting bail hearings. In determining bail in Smith's case, the judge considers all of the relevant factors, including the severity of the crime, the strength of the evidence, the defendant's prior criminal record, and his background (such as his employment history, family history, and other personal information). The judge ultimately determines that Smith is a high risk for flight. The judge sets bail at $20,000.

C. Smith is now in front of a judge at a hearing to determine whether bail should be set for this most recent arrest. According to state law, bail should be set at the lowest possible level that is necessary to ensure that the defendant returns to court for trial. There are no bail bondsmen in this jurisdiction; thus, if bail is set, Smith or his family would have to pay the entire amount of whatever bail is set. This court uses a computer program at bail hearings to determine how likely the defendant is to return to court. The computer program considers all of the relevant factors, including the severity of the crime, the strength of the evidence, the defendant's prior criminal record, and his background (such as his employment history, family history, and other personal information) The computer program is able to reference data from hundreds of thousands of cases from across the state in which defendants were released pending trial and it compares all of those cases to the defendant's case to determine his likelihood of failing to appear in court for trial. The computer program has a high degree of accuracy in predicting which defendants are likely to appear in court and which are likely to not appear in court. In Smith's case, the computer program determines that Smith is not a high risk for flight because 90% of people with his characteristics and in his situation return to court for trial. The judge accepts the program's assessment and sets bail at $4,000.

D. Smith is now in front of a judge at a hearing to determine whether bail should be set for this most recent arrest. According to state law, bail should be set at the lowest possible level that is necessary to ensure that the defendant returns to court for trial. There

are no bail bondsmen in this jurisdiction; thus, if bail is set, Smith or his family would have to pay the entire amount of whatever bail is set. This court uses a computer program at bail hearings to determine how likely the defendant is to return to court. The computer program considers all of the relevant factors, including the severity of the crime, the strength of the evidence, the defendant's prior criminal record, and his background (such as his employment history, family history, and other personal information) The computer program is able to reference data from hundreds of thousands of cases from across the state in which defendants were released pending trial and it compares all of those cases to the defendant's case to determine his likelihood of failing to appear in court for trial. The computer program has a high degree of accuracy in predicting which defendants are likely to appear in court and which are likely to not appear in court. In Smith's case, the computer program determines that Smith is a high risk for flight because 90% of people with his characteristics and in his situation do not return to court for trial. The judge accepts the program's assessment and sets bail at $20,000.

[All respondents then were asked the following questions:]

What is your opinion of the amount of bail that was set in this case?

- Far too low
- Too low
- About right
- Too high
- Far too high
- No opinion

What amount of bail would you set if you were the judge?

_____

Do you think that Smith was treated fairly in the bail hearing?

- Yes, the process was very fair to Smith
- Yes, the process was fair to Smith
- The process seemed neither particularly fair nor unfair to Smith
- No, the process was unfair to Smith

- No, the process was very unfair to Smith
- No opinion

Why do you think the process was fair or unfair?

_____

APPENDIX B: COMMENTS FROM RESPONDENTS

All comments that mention the computer program from the "machine/ low bail" respondents who thought the process was fair:

- It wasn't really a human judgement.
- Based on Smith's history and software program was more than fair.
- I think it was fair because the program was based on characteristics that would influence their likelihood to return to court
- Because the judge used a computer program and there couldn't have been bias.
- It was fair because it evaluated by a machine rather than a person.
- Bail was set the lowest possible level for Smith based on computer (no emotion involved). Smith has previous criminal records and I thought the way he was treated was fair.
- The methods used to make the decision were fair - it seems like an unbiased algorithm made the decision. If anything, the outcome was kind of lenient in the defendant's favor, so I think it was a fair outcome/process.
- The computer used data to determine the risks.
- That an algorithm was used makes it pretty fair to begin with, basing the decision on factors known to be relevant. No personal bias had room to interfere.
- A computer that is highly accurate was used to determine the bail. The bail was very low in my opinion thus he was treated overly fair.
- Did not consider him a flight risk and set low bail regardless of whether computer program and judge used adequate analysis.
- The bail was set according to a computer program or algorithm and not via the judge's thoughts, which could be affected by emotion.

- it was fair because a computer program issued it
- While the process was fair in that it took the judgement out of the hands of humans so there could be no bias, the computer program seems to have been unfair to society in setting the bail so low.
- It was more fair to be processed by a computer program than by letting the bias of a human affect judgement of his flight risk
- It was only fair due to the program stating such a low bail amount in favor of him.
- It was fair as the computer program set the bail and no other factors were brought in such as race, age, or gender.
- It followed the rules and regulations and adhered to computer logic.
- It was the product of an algorithm. It was based on his past actions
- The computer program is unbiased

All comments that mention the computer program from the "machine/ high bail" respondents who thought the process was fair:

- I think the program was fair as this individual has already missed one court date and has committed a violent crime.
- The program seems to have efficient algorithm and can make good decisions.
- A computer cannot be biased in any way because it's just a machine.
- Despite being a flight risk and the severity of his crime, his bail was set fairly low. I think that using a computer program to assess flight risk is fair.
- It's statistical algorithm based on history. he has failed to appear in court before. He is on a third strike, A firearm was used.

- Since the assessment is done by computer logic, it seems about fair in judging the probability of Smith committing the crime again
- A computer was used which would have no bias.
- it was objective, free from human bias for the most part, and a direct assessment based on Smith as an individual, therefore, the process was quite fair to Smith.
- I think the judge and the program accurately assessed the nature of the crime and Mr. Smith's likelihood of fleeing.
- I think it's fair because the judge made his determination based on the program, which compares John's type of cases, etc. According to the information above the program is very accurate, so I think it was a fair assessment.
- He has previous issues with showing up to court, he was in possession of the stolen items and they analyzed similar individuals on the computer program.
- I think the factors that were taken into account were appropriate, and presumably the algorithm is taking into account much of the thinking that would otherwise go into setting his bail. In fact, it's probably a better predictor of his behavior than a human thinking about it. Besides, his past history suggests he should get no favors.
- It seems fair because the computer has no prejudices and computes the estimate based on facts. The bail amount however seems unreasonably high.
- A computer program was used to determine bail and not strictly the judgment of the judge.
- a computer program determines the amount based on all info available.
- Since the bail amount was determined through a computer program, human bias was at a minimum in this case.
- It was based on a lot of factors and the program appears to be fairly accurate at 90%. I think that is about as fair as one can be.

- It was fair and impartial because a computer analyzed and set the bail. There was no human bias.
- I feel like that using a machine would be the most impartial and fair way to determine something like this. No opinions, just the presented facts.
- Although I think The algorithm is fair the bail was just way too high for the crime
- The process was fair because it used a computer to make comparisons and calculations and didn't just go on the opinion of the judge alone
- it was set by a computer and didn't seem to have biases like a human would. also he's clearly going to keep committing crimes.
- I think it is fair because it says that the computer program has a high degree of accuracy in cases like this and so I trust the computer program because it is objective and unbiased.
- A computer with the sorts of parameters that were explained in the text can be more fair than a person who may have biases or prejudice influence their decision.

All comments that mention the computer program from the "machine/ low bail" respondents who thought the process was unfair:

- The decisions were made by a computer.
- it uses a computer to determine this which is not as fair as a real judge
- Because they used a computer program to make the decision. An actual person should do this. Computers aren't always right
- Computer generated may not take into account extenuating circumstances.
- The quality of the computer software is in question. I don't know that it's trustable software. However, his bond was set at a reasonable rate for the crimes alleged.
- It's use of a program is not very good.

All comments that mention the computer program from the "machine/ high bail" respondents who thought the process was unfair:

- I think the judge has to think for himself instead of using a program with no human thought connected to it.
- I think it is somewhat fair in that computer algorithms are never 100% accurate.
- A computer program should not determine someone's bail amount - that is why we have judges in courtrooms and pretrial booking officer's to determine that measure.
- how can we believe the computer assessments to a victim? it is totally unfair, he do not deserves a bail.
- computer based decision on these cases are unreliable
- I don't believe computers will ever be able to replace the judgement of a human being.
- because it doesn't take his personal ability to pay into account and his ability to show should not be based on a computer program
- I think the bail amount was more than fair for armed robbery but I'm still a little hesitant to be letting a computer determine things like that.