2017

Constructing Recidivism Risk

Jessica M. Eaglin
Indiana University Maurer School of Law, jeaglin@indiana.edu

Follow this and additional works at: https://www.repository.law.indiana.edu/facpub

Part of the Criminal Law Commons, and the Criminology and Criminal Justice Commons

LAW LIBRARY
INDIANA UNIVERSITY
Maurer School of Law
Bloomington

CONSTRUCTING RECIDIVISM RISK

Jessica M. Eaglin[*]

ABSTRACT

Courts increasingly use actuarial—meaning statistically derived—information about a defendant's likelihood of engaging in criminal behavior in the future at sentencing. This Article examines how developers construct the tools that predict recidivism risk. It exposes the numerous choices that developers make during tool construction with serious consequences to sentencing law and policy. These design decisions require normative judgments concerning accuracy, equality, and the purpose of punishment. Whether and how to address these concerns reflects societal values about the administration of criminal justice more broadly. Currently, developers make these choices in the absence of law, even as they face distinct interests that diverge from the public. As a result, the information produced by these tools threatens core values at sentencing. This Article calls for accountability measures at various stages in the development process to ensure that the resulting risk estimates reflect the values of the jurisdictions where the tools will be applied at sentencing.

INTRODUCTION

Predictive technologies increasingly appear at every stage of the criminal justice process.[1] From predictive policing to pretrial bail to sentencing, public and private entities outside the justice system now construct policy-laden evidence of recidivism risk to facilitate the administration of justice.[2] Using the actuarial risk tools for sentencing as illustration, this Article examines the normative judgments entailed in the development of predictive recidivism risk information for the administration of justice. It proposes measures to infuse public input into tool construction.

Criminal courts increasingly engage in "risk-based sentencing" as states consider and adopt more data-driven criminal justice reforms.[3] Risk-based sentencing occurs when a court relies on actuarial risk assessment tools that predict a defendant's likelihood of engaging in criminal behavior to inform and

---

[1] Predictive technologies are spreading through the criminal justice system like wildfire. *See, e.g.*, Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327 (2015) (explaining predictive policing and Fourth Amendment reasonable suspicion determinations); Cecelia Klingele, *The Promises and Perils of Evidence-Based Corrections*, 91 NOTRE DAME L. REV. 537, 564–67 (2015) (explaining risk assessments for probation and parole hearings); Sandra G. Mayson, *Bail Reform and Restraint for Dangerousness: Are Defendants a Special Case?*, 127 YALE L.J. (forthcoming 2017) (discussing risk assessments at pretrial bail hearings); Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871 (2016) (explaining program-predicted criminal activity and Fourth Amendment reasonable suspicion determinations); Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803 (2014) (discussing risk assessments at sentencing). *See generally* BERNARD E. HARCOURT, AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE (2007) (discussing dilemmas of prediction in various stages of criminal process).

[2] *See, e.g.*, Julia Angwin et al., *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (describing leading risk assessment tools for sentencing and corrections developed by Northpointe); Ellen Huet, *Server and Protect: Predictive Policing Firm PredPol Promises to Map Crime Before It Happens*, FORBES (Feb. 11, 2015, 6:00 AM), https://www.forbes.com/sites/ellenhuet/2015/02/11/predpol-predictive-policing (discussing the leading predictive policing software PredPol); *Public Safety Assessment: Risk Factors and Formula*, LAURA AND JOHN ARNOLD FOUNDATION (2016), http://www.arnoldfoundation.org/wp-content/uploads/PSA-Risk-Factors-and-Formula.pdf (describing a risk tool for pretrial bail hearings developed by a nonprofit foundation).

[3] *See, e.g.*, Starr, *supra* note 1; Anna Maria Barry-Jester et al., *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. Scholars and policymakers often refer to this practice as "evidence-based" sentencing because it is part of a larger shift towards "evidence-based" practices in criminal justice. *See* Klingele, *supra* note 1. This Article will not use that phrase because it is misleading in this context, as courts already use evidence to determine a sentence. *See infra* Part I. This practice is new in the sense that courts use actuarial risk information. Thus, this Article refers to the practice as "risk-based sentencing." *Cf.* Melissa Hamilton, *Adventures in Risk: Predicting Violent and Sexual Recidivism in Sentencing Law*, 47 ARIZ. ST. L.J. 1 (2015).

guide its discretion in sentencing.[4] These actuarial—meaning statistically derived—tools assess individuals based on a series of factors to produce a score that ranks defendants according to likelihood of engaging in specified behavior in the future.[5] Judges may consider the information provided by recidivism risk tools directly in the sentencing process, or probation officers may confront the tools and collapse the information into a presentence recommendation to the court.[6] This information may influence any number of sentencing determinations, including whether to impose probation versus incarceration, the length of incarceration, and the types of conditions a judge may impose on probation.[7]

A growing body of scholarship considers the entry of risk-based sentencing practices in the states. Scholars debate the use of actuarial risk information at sentencing for very different reasons. Advocates contend that, because risk tools more objectively and consistently predict the likelihood of recidivism than the inevitable human guesswork of judges,[8] using the tools at sentencing will improve accuracy.[9] More accuracy, they suggest, will improve sentencing practices.[10] Critics oppose risk-based sentencing as a matter of fairness. They contend that, because risk tools rely on factors like gender or proxies for race, using the tools at sentencing is impermissible as a matter of constitutionality or

---

[4] John Monahan & Jennifer L. Skeem, *Risk Redux: The Resurgence of Risk Assessment in Criminal Sanctioning*, 26 FED. SENT'G REP. 158, 159 (2014); Starr, *supra* note 1, at 805.

[5] *See, e.g.*, John Monahan, *A Jurisprudence of Risk Assessment: Forecasting Harm Among Prisoners, Predators, and Patients*, 92 VA. L. REV. 391, 405–06 (2006).

[6] Monahan & Skeem, *supra* note 4, at 159.

[7] *See* John Monahan & Jennifer L. Skeem, *Risk Assessment in Criminal Sentencing*, 12 ANN. REV. CLINICAL PSYCHOL. 489, 493–94 (2016) (discussing length of sentence, diversion, and interventions); PAMELA M. CASEY ET AL., NAT'L CTR. FOR STATE COURTS, USING OFFENDER RISK AND NEEDS ASSESSMENT INFORMATION AT SENTENCING: GUIDANCE FOR COURTS FROM A NATIONAL WORKING GROUP 8–10 (2011), http://www.ncsc.org/~/media/Microsites/Files/CSI/RNA%20Guide%20Final.ashx (focusing on diversion from prison to probation).

[8] Johnson v. United States, 135 S. Ct. 2551, 2557–58 (2015) (discussing the "judicial assessment of risk"); Monahan, *supra* note 5, at 427–28 (discussing the accuracy of such tools).

[9] *See, e.g.*, Jordan Hyatt et al., *Reform in Motion: The Promise and Perils of Incorporating Risk Assessments and Cost-Benefit Analysis into Pennsylvania Sentencing*, 49 DUQ. L. REV. 707, 713 (2011) ("The ability to generate accurate assessments that can be systematically used in the sentencing courtroom will represent an improvement over current practices.").

[10] *See, e.g.*, NATHAN JAMES, CONG. RESEARCH SERV., RISK AND NEEDS ASSESSMENT IN THE CRIMINAL JUSTICE SYSTEM 1 (2015) ("Assessment instruments might help increase the efficiency of the justice system by identifying low-risk offenders who could be effectively managed on probation rather than incarcerated, and they might help identify high-risk offenders who would gain the most by being placed in rehabilitative programs.").

bad policy.[11] This scholarship influences larger debates about whether and how to incorporate predictive risk information into the administration of justice.[12] Yet none of these scholars consider how to regulate the production of risk information. Instead, they debate whether to eliminate its use entirely.

Outside the sentencing context, a growing body of scholarship examines the rise of predictive analytics used both within the criminal justice system[13] and outside of it.[14] These scholars largely call for accountability measures that ensure predictions are consistent with normative concepts of fairness.[15] Yet few of these scholars engage with the underlying normative debates implicit in the construction of the tools. Few urge elimination of the tools all together.[16]

This Article enters at the intersection of these two bodies of scholarship. It exposes how external incentives intersect with law and policy in the construction of risk tools for sentencing. How tools are constructed has great import to the "truths" the resulting outcomes purport to assert. Using actuarial risk tools used for sentencing as illustration, this Article does two things. First, it systematically exposes the normative judgments embedded in actuarial risk assessment tools' construction. Second, it calls for legal accountability to

---

[11] See HARCOURT, supra note 1, at 3–6. For constitutional debate, compare J.C. Oleson, *Risk in Sentencing: Constitutionally Suspect Variables and Evidence-Based Sentencing*, 64 SMU L. REV. 1329 (2011) (arguing that risk-based sentencing practices are constitutional), with Dawinder S. Sidhu, *Moneyball Sentencing*, 56 B.C. L. REV. 671 (2015) (arguing that risk-based sentencing practices are unconstitutional), and Starr, *supra* note 1 (arguing that risk-based sentencing is unconstitutional). For normative debate, compare Hyatt et al., *supra* note 9 (arguing that using risk-based sentencing practices instills fairness into the criminal justice process), with Hamilton, *supra* note 3 (arguing that risk-based sentencing practices are prejudicial and unreliable), and Bernard E. Harcourt, *Risk as a Proxy for Race: The Dangers of Risk Assessment*, 27 FED. SENT'G REP. 237 (2015) (arguing that risk-assessment tools aggravate racial disparity in the criminal justice system).

[12] See Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. (forthcoming 2017).

[13] See, e.g., ERIN E. MURPHY, INSIDE THE CELL: THE DARK SIDE OF FORENSIC DNA (2015) (examining the risks of DNA testing used in criminal trials); Ferguson, *supra* note 12 (discussing predictive technologies and realities unique to the criminal justice system); Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CALIF. L. REV. 721, 723 (2007) (discussing new forensic techniques introduced at various stages of the criminal justice process).

[14] See, e.g., Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671 (2016) (discussing unintended discriminatory effects of data mining); Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857 (2017) (describing the use of data analytic tools in the workplace).

[15] See, e.g., FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2015); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 6 (2014).

[16] But see HARCOURT, *supra* note 1; Harcourt, *supra* note 11.

ensure risk-tool construction in service of the law.[17] Actuarial risk assessment tools obscure difficult normative choices about the administration of criminal justice. This Article proposes a framework to pierce the opacity of these tools with various interventions to facilitate public discourse and input throughout the construction process.

Entities developing actuarial risk assessment tools for sentencing make policy assumptions during construction that relate to highly contested and undecided questions of sentencing law and policy. Part I unpacks the tool-construction process to demonstrate what decisions tool developers make and when. It divides this process into two stages, each of which implicates normative judgments about sentencing law and policy in different ways. During the first stage, researchers decide what data to collect, where to collect data from, how to define recidivism, and what predictive factors to observe in the data set.[18] They also create an algorithm to reflect their conclusions about recidivism risk.[19] These decisions tie into legal questions about what counts at sentencing and how these factors should be weighted.

The second stage occurs when entities decide how to convey the algorithm's results for use by criminal justice actors.[20] Public and private entities translate the algorithmic outcomes into recidivism risk categories. These decisions implicate policy questions about who should be considered a risk and how much risk society tolerates. Combined, this examination demonstrates that actuarial risk tools, while "scientific" in the sense that developers use technology to assess risk, reflect normative judgments familiar to sentencing law and policy debates.[21] Yet, unlike previous efforts to infuse prediction into sentencing, it is difficult to identify the normative judgments reflected in the information produced by the tools.

---

[17] *See, e.g.,* Sheila Jasanoff, *Serviceable Truths: Science for Action in Law and Policy,* 93 TEX. L. REV. 1723, 1730 (2015) (calling for a shift from inquiries about validation of scientific claims to a more normative concept of "serviceable truth"). Understanding risk technology as a "serviceable truth" requires striking the balance between "scientific facts and reasons on the one hand and the nurture and protection of human lives and flourishing on the other," and recognizing that "science's role in the legal process is not simply, even preeminently, to provide a mirror of nature. Rather it is to be of service to those who come to the law with justice or welfare claims whose resolution happens to call for scientific fact-finding." *Id.* (emphasis omitted).

[18] *See infra* Sections I.A.1–3.

[19] *See infra* Section I.A.4.

[20] *See infra* Section I.B.

[21] *See infra* Part I; *see also* Erica Beecher-Monas & Edgar Garcia-Rill, *Danger at the Edge of Chaos: Predicting Violent Behavior in a Post-*Daubert *World,* 24 CARDOZO L. REV. 1845, 1896 (2003) (stating that "risk is a social construct" and not an exact science).

Part II explores the significance of construction choices with respect to three normative and deeply contested societal values central to sentencing law and policy and the administration of criminal justice more broadly. Section II.A considers tool construction and the notion of accuracy. Scientific studies may demonstrate that a tool is "accurate" because it differentiates between defendants who do or do not engage in specific behavior in the future more consistently than chance or it classifies a defendant who engages in particular behavior in the future as high risk versus identifying a defendant who commits no future crime as low risk.[22] Both of these types of accuracy relate to the overarching aims of the justice system, but neither assessment provides insight as to whether a tool credibly meets those aims. Entities developing risk tools cannot answer these questions through empirical assessment; only society can make that determination.[23]

Section II.B considers how tool-construction choices compromise equality at sentencing. Risk tools inevitably classify defendants from historically disadvantaged backgrounds—particularly black men—as higher risk than other defendants due to construction decisions.[24] As a result, certain defendants will not have equal opportunity to benefits that may flow from the introduction of risk tools at sentencing. Whether and how much to compromise this value is a matter that society should address before tool adoption. The final value considered in section II.C relates to the purpose of punishment. Information about a defendant's likelihood of engaging in criminal behavior in the future may further utilitarian purposes of punishment depending on how it is developed.[25] Society should decide whether and how to incorporate this information into sentencing practices.[26]

The exploration in sections II.A–C demonstrates that the construction of a recidivism risk tool is coterminous with the law's normative concerns and institutional practices. Yet the entities developing risk tools often decide these difficult questions without guidance from law or policymakers. This is true despite the unique set of interests that tool developers' face when constructing

---

[22] See Hamilton, *supra* note 3, at 24 (explaining that risk tool accuracy is often represented through predictive validity studies).

[23] See *infra* Section II.A.

[24] See *infra* Section II.B.

[25] See, e.g., Christopher Slobogin, *The Civilization of the Criminal Law*, 58 VAND. L. REV. 121 (2005) (urging consequentialism over retributivism as the guiding purpose of punishment at sentencing).

[26] See *infra* Section II.C.

a risk tool.[27] Section II.D also demonstrates how the desire for cheap, varied, and easily accessible data incentivizes developers to make construction choices that may contradict or conflict with a state's existing sentencing policies and practices.

Part III provides a path forward. It calls for democratic engagement with the construction of actuarial risk tools. Whether and how a risk assessment tool predicts recidivism in the administration of criminal justice requires accountability to the normative values of the community where a tool is applied. While scholars largely discuss accountability in scientific terms,[28] section III.A explores a separate meaning of the term that builds from the realities of applying the tool's outcomes at sentencing. Accountability in this context requires removing the veil of objectivity to facilitate community engagement with the normative judgments underlying tool construction.[29] Section III.B calls on tool developers and government actors to facilitate this democratic accountability in the construction of risk. It identifies three levels of opacity that prevent meaningful engagement, and suggests various interventions to infuse criminal justice expertise and political process accountability into the tool-construction process. These reforms heed the essence of calls for caution in automated systems; namely, that tools should reflect societal values and ensure democratic input in construction.[30]

This Article provides two novel contributions to existing literature. First, it sharpens theoretical critiques about using risk tools at sentencing and broadens the scope of the ongoing normative debate about whether states should adopt risk-based sentencing practices.[31] Although the public is exposed to the scholarly debate on risk-based sentencing,[32] its voice is largely absent

---

[27] See infra Section II.D.

[28] See, e.g., Joshua A. Kroll et al., Accountable Algorithms, 165 U. PA. L. REV. 633 (2017) (discussing computer science accountability). But see Anupam Chander, The Racist Algorithm?, 115 MICH. L. REV. 1023 (2017) (discussing accountability as a matter of both democratic and computer science significance).

[29] As Sheila Jasanoff explains, "objectivity itself is better understood not as an intrinsic attribute of science but as a perceived characteristic of scientific knowledge, arrived at through culturally conditioned practices." Jasanoff, supra note 17, at 1739–40. Similarly, the perceived objectivity of technology used to produce recidivism risk knowledge for sentencing is constructed.

[30] See Chander, supra note 28; Danielle Keats Citron, Technological Due Process, 85 WASH. U. L. REV. 1249, 1258 (2008); Citron & Pasquale, supra note 15, at 18–20; Kroll et al., supra note 28.

[31] See, e.g., HARCOURT, supra note 1; Jessica M. Eaglin, Against Neorehabilitation, 66 SMU L. REV. 189 (2013); Hamilton, supra note 3; Harcourt, supra note 11; Starr, supra note 1; Michael Tonry, Legal and Ethical Issues in the Prediction of Recidivism, 26 FED. SENT'G REP. 167, 167 (2014).

[32] See Eric Holder, Attorney Gen., Remarks at the National Association of Criminal Defense Lawyers 57th Annual Meeting and 13th State Criminal Justice Network Conference (Aug. 1, 2014),

regarding the challenges that stem from tool construction. Second, this Article joins a growing body of scholarship on the risks of applying big data techniques to the administration of criminal justice. Whether predictive analytics produce evidence that should be relied upon in the criminal justice system is an apparent yet under-theorized component to this development.[33] This Article lays foundation for the expansion of that discourse by explaining why caution and accountability measures are necessary when predicting recidivism risk.

I. CONSTRUCTING RECIDIVISM RISK TOOLS

While recidivism risk has long influenced criminal justice outcomes, the use of actuarial tools heralds a new, data-centric approach to prediction in sentencing. Initial attempts to use prediction in sentencing determinations relied upon clinical assessment of recidivism risk to inform parole release decisions.[34] By the 1980s, states and the federal system began to introduce the "science of probabilities" into sentencing law through a variety of methods. For example, sentencing commissions incorporated criminal history into determinate sentencing guidelines as a measure of recidivism risk.[35] Specific

---

https://www.justice.gov/opa/speech/attorney-general-eric-holder-speaks-national-association-criminal-defense-lawyers-57th; Sonja B. Starr, Opinion, *Sentencing, by the Numbers*, N.Y. TIMES (Aug. 10, 2014), https://www.nytimes.com/2014/08/11/opinion/sentencing-by-the-numbers.html.

[33] For more general insight on this discourse in the context of trials, see, for example, Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972 (2017) [hereinafter Roth, *Machine Testimony*]; Andrea Roth, *Trial by Machine*, 104 GEO. L.J. 1245 (2016) [hereinafter Roth, *Trial by Machine*]; Rebecca Wexler, *Life, Liberty and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. (forthcoming 2018). *See also* Erin Murphy, *The Mismatch Between Twenty-First-Century Forensic Evidence and Our Antiquated Criminal Justice System*, 87 S. CAL. L. REV. 633 (2014) (discussing the failure of the criminal justice system to handle high-tech evidence); Murphy, *supra* note 13 (explaining the use of DNA typing, data mining, location tracking, and biometric technologies).

[34] Under the indeterminate sentencing structures prevalent until the late 1970s, parole boards frequently used clinical assessments of recidivism risk to inform choices about whether and when to release an offender on parole. *See* HARCOURT, *supra* note 1, at 52–55. These assessments were "clinical" in the sense that professional psychologists interviewed defendants, asking a series of unguided questions to determine whether the defendant would commit a crime in the future. *See id.* at 40–42. The expert "relied on whatever information the individual clinician deemed pertinent" to produce a recidivism risk prediction. Christopher Slobogin, *Dangerousness and Expertise Redux*, 56 EMORY L.J. 275, 283 (2006); *see also* Barbara D. Underwood, *Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment*, 88 YALE L.J. 1408, 1423 (1979) ("A clinical decisionmaker is not committed in advance of decision to the factors that will be considered and the rule for combining them.").

[35] Professor Paul Robinson, a former commissioner on the U.S. Sentencing Commission, explained in 2001, "[t]he rationale for heavy reliance upon criminal history in sentencing guidelines is its effectiveness in incapacitating dangerous offenders." Paul H. Robinson, Commentary, *Punishing Dangerousness: Cloaking Preventive Detention as Criminal Justice*, 114 HARV. L. REV. 1429, 1431 n.7 (2001). State sentencing

calls to use selective incapacitation theory at sentencing also reflected an interest in prediction.[36] This theory supported the proliferation of legislatively created sentencing reforms that relied predominately on prior criminal history, including mandatory minimum penalties and "three strikes" laws.[37] These sentencing reforms enhanced punishment for prior offenders because studies indicated that prior criminal history was the best predictor of recidivism.[38] Most reforms introduced enhancements by aggregating similar crimes or defendants with similar criminal histories so that judges would increase sentence length in certain types of cases.[39]

The recent resurgence in prediction at sentencing is different. Entities inside and outside the justice system now produce risk assessment tools that estimate an individual's likelihood of recidivism based on factors not necessarily connected to the criminal justice system. The tools assess recidivism risk based on actuarial—or statistical—analysis of data-driven observations about previously arrested or convicted individuals' past behavior.[40] The tools rank and classify a defendant based on a series of identified factors that correlate with the occurrence of specific criminalized behavior.[41] Many of these factors do not relate to the defendant's offense of

---

guidelines likely use similar logic to support development of guideline systems that rely predominately on prior criminal history as well. See HARCOURT, *supra* note 1, at 91–92; *see also* RICHARD S. FRASE ET AL., ROBINA INST. OF CRIMINAL LAW & CRIMINAL JUSTICE, CRIMINAL HISTORY ENHANCEMENTS SOURCEBOOK 14–16 tbl.1.1 (2015), https://robinainstitute.umn.edu/publications/criminal-history-enhancements-sourcebook (identifying at least five states that explicitly justify criminal history enhancement based on risk, but noting that the majority of states do not explain why they enhance sentences based on prior criminal history).

[36] Selective incapacitation refers to a theory of punishment focused on predicting the offenders capable of rehabilitation and those who have a high risk of reoffending and should thus be incapacitated for extended terms. See Eaglin, *supra* note 31, at 222–23.

[37] HARCOURT, *supra* note 1, at 91–93. For a more detailed description of these laws, see Jessica M. Eaglin, *The Drug Court Paradigm*, 53 AM. CRIM. L. REV. 595, 601, 615–16 (2016).

[38] HARCOURT, *supra* note 1, at 91.

[39] Professor Albert Alschuler recognized that the sentencing guidelines reflected a "changed attitude towards sentencing" that emphasizes "rough aggregations and statistical averages" about "collections of cases and . . . social harm," rather than "individual offenders and the . . . circumstances of their cases." Albert W. Alschuler, *The Failure of Sentencing Guidelines: A Plea for Less Aggregation*, 58 U. CHI. L. REV. 901, 951 (1991).

It is worth noting that consideration of *individual* risk at sentencing declined by the 1990s as states shifted focus toward reducing unwarranted disparities and imposing retributive punishment. Although considerations of risk remained when determining treatment interventions, its use to determine the nature or duration of a sentence became highly suspect. See Tonry, *supra* note 31, at 167. This method of prediction is now experiencing a resurgence. See id.

[40] See HARCOURT, *supra* note 1, at 1–2.

[41] There are two types of risk assessment tools—those that pre-identify risk factors ("checklist tools") and those that allow the computer to derive predictive factors ("machine learning tools"). See RICHARD BERK,

conviction or criminal history. When used at sentencing, the tool's outcome estimates the likelihood of a defendant engaging in criminal behavior in the future.[42] That result can influence determinations about whether the defendant should be diverted from incarceration to probation, the length of criminal justice supervision, or specific interventions made available during that term of supervision.[43]

A variety of public and private entities currently develop recidivism risk assessment instruments. Commercial risk tools occupy a large space in the field of prediction at sentencing.[44] Private companies have developed some of the leading tools used in several jurisdictions. These include the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) assessment tool designed by Northpointe Institute for Public Management, Inc.[45] and the Level of Service Inventory-Revised (LSI-R) designed by Multi-Health Systems, Inc.[46]

CRIMINAL JUSTICE FORECASTS OF RISK: A MACHINE LEARNING APPROACH 18 (2012) (describing simple cross-tabulation tools versus complex data mining tools); *see also* Mayson, *supra* note 1, at 9–11 (distinguishing between "checklist" and "machine forecasting" tools). The most prevalent risk tools used at sentencing are checklist tools. *See, e.g.*, PAMELA M. CASEY ET AL., NAT'L CTR. FOR STATE COURTS, OFFENDER RISK & NEEDS ASSESSMENT INSTRUMENTS: A PRIMER FOR COURTS app. at A-31 (2014), http://www.ncsc.org/~/media/Microsites/Files/CSI/BJA%20RNA%20Final%20Report_Combined%20Files%208-22-14.ashx; JAMES, *supra* note 10, at tbl.B-1 (canvasing leading risk and needs instruments). As such, these tools are the focus of this Article. *See infra* notes 51–52, 56–58. However, researchers are steering risk tool development in the direction of machine learning. *See* Richard Berk & Jordan Hyatt, *Machine Learning Forecasts of Risk to Inform Sentencing Decisions*, 27 FED. SENT'G REP. 222 (2015). There, the computer identifies factors to estimate risk based on constantly updated data and more complex and powerful algorithms. *See, e.g.*, Kroll et al., *supra* note 28, at 638. Such tools will present unique challenges at sentencing, some of which are addressed here. *See infra* Section III.B.3. More nuanced research promises to address the specific challenges these tools present in more detail in the future.

   [42] This development is consistent with the shift towards a new penology of punishment described by Professors Malcolm Feeley and Jonathan Simon in the early 1990s. *See* Malcolm M. Feeley & Jonathan Simon, *The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications*, 30 CRIMINOLOGY 449, 455 (1992) ("The new penology . . . is about identifying and managing unruly groups."). As noted elsewhere, this approach is crystallized in neorehabilitative reforms, including the use of actuarial risk tools at sentencing. *See generally* Eaglin, *supra* note 31. This Article expands on the previous observations of Professors Feeley and Simon by examining one of the "new technologies to identify and classify risk" highlighted in their previous work. *See* Feeley & Simon, *supra*, at 457.

   [43] *See supra* note 7.

   [44] Monahan & Skeem, *supra* note 7, at 499 (recognizing the wide array of "[c]ommercial off-the-shelf tools" developing for use in sentencing alongside government designed tools). This is consistent with the broader reality that private sector industries develop, market, and maintain most technology devices and tools used in the criminal justice system, including GPS tracking devices, biometrics, and the like. Erin Murphy, *The Politics of Privacy in the Criminal Justice System: Information Disclosure, the Fourth Amendment, and Statutory Law Enforcement Exemptions*, 111 MICH. L. REV. 485, 536 (2013).

   [45] NORTHPOINTE, INC., PRACTITIONER'S GUIDE TO COMPAS CORE (2015), https://assets.documentcloud.org/documents/2840784/Practitioner-s-Guide-to-COMPAS-Core.pdf (discussing Northpointe's risk scales for

Some state sentencing commissions have developed state-specific predictive tools for sentencing. For example, the Virginia Sentencing Commission created several risk assessment tools used at sentencing in the state.[47] These include a nonviolent risk assessment tool used to encourage non-prison sanctions for the lowest risk defendants and a separate predictive tool applied to sex offenders.[48] The Pennsylvania Commission on Sentencing is in the final stages of developing its own risk assessment tool for sentencing, too.[49]

In addition, other public-private entities are developing risk tools. For example, Canadian psychologists and professors developed the popular Violence Risk Appraisal Guide (VRAG).[50] Statisticians at the University of Cincinnati Corrections Institute (UCCI) developed the Ohio Risk Assessment System (ORAS), which includes several tools used at sentencing.[51] Private non-profit organizations also fund independent development of risk prediction

---

general recidivism, violent recidivism, and pretrial misconduct). Please note that Northpointe, Inc., recently rebranded itself as equivant. All product lines remain intact, including COMPAS. *See Courtview, Constellation & Northpointe Re-brand to Equivant,* EQUIVANT, http://www.equivant.com/blog/we-have-rebranded-to-equivant.

[46] CASEY ET AL., *supra* note 41, app. at A-38.

[47] Richard P. Kern & Meredith Farrar-Owens, *Sentencing Guidelines with Integrated Offender Risk Assessment,* 25 FED. SENT'G REP. 176 (2013).

[48] *Id.* at 177.

[49] *See* 42 PA. STAT. AND CONS. STAT. ANN. § 2154.7 (West Supp. 2017) (requiring the commission to develop a risk assessment instrument for sentencing); *Risk Assessment Project,* PA. COMMISSION ON SENT'G (2017), http://pcs.la.psu.edu/publications-and-research/research-and-evaluation-reports/risk-assessment.

[50] *See* GRANT T. HARRIS ET AL., VIOLENT OFFENDERS: APPRAISING AND MANAGING RISK 5–7 (3d ed. 2015).

[51] *Risk Assessment,* UNIV. OF CINCINNATI CORR. INST. (2017), http://cech.uc.edu/centers/ucci/services/risk-assessment.html (describing risk assessment tools developed for Ohio). *See generally* CASEY ET AL., *supra* note 41, at app. A-52–56 (explaining the use of ORAS at sentencing and its development). UCCI replicated the ORAS system for use in Indiana. *The Indiana Risk Assessment System (IRAS) and the Indiana Youth Assessment System (IYAS),* IND. JUDICIAL BRANCH, http://www.in.gov/judiciary/cadp/2762.htm. The Indiana Risk Assessment System (IRAS), based on the ORAS system, similarly does not indicate that it was specifically designed for post-conviction sentencing purposes. *See, e.g.,* PAMELA M. CASEY ET AL., NAT'L CTR. FOR STATE COURTS, USE OF RISK AND NEEDS ASSESSMENT INFORMATION AT SENTENCING: GRANT COUNTY, INDIANA 5 (2013), http://www.ncsc.org/~/media/Microsites/Files/CSI/RNA%20Brief%20-%20Grant %20County%20IN%20csi.ashx. Rather, UCCI validated the tools used in Ohio, including a pretrial tool, a community supervision tool, and a reentry tool for use in the state. Indiana sentencing courts are encouraged to use complementary risk tools (often commercial) to supplement the IRAS system. *See* UNIV. OF CINCINNATI, INDIANA RISK ASSESSMENT SYSTEM i–iii (2010), http://www.pretrial.org/download/risk-assessment/Indiana %20Risk%20Assessment%20System%20(April%202010).pdf (discussing use of the tool).

tools. The Laura and John Arnold Foundation and the MacArthur Foundation lead in these efforts.[52]

Some tools predict recidivism generally, while others estimate recidivism for specific types of offenses.[53] Certain tools purport to predict whether a person is at risk of committing a violent crime,[54] a sex offense,[55] or some other specific type of offense.[56] Tools may be designed for distinct criminal justice use. Certain tools, like the VRAG and the Static-99, use limited, immutable

---

[52] The Laura and John Arnold Foundation focuses on developing risk assessment tools for use in pretrial bail determinations. *See Developing a National Model for Pretrial Risk Assessment*, LAURA & JOHN ARNOLD FOUND. (Nov. 2013), http://www.arnoldfoundation.org/wp-content/uploads/2014/02/LJAF-research-summary_PSA-Court_4_1.pdf. The MacArthur Foundation played a critical role in development of risk tools for use in mental health services. JOHN MONAHAN ET AL., RETHINKING RISK ASSESSMENT: THE MACARTHUR STUDY OF MENTAL DISORDER AND VIOLENCE (2001). Both are looking to expand their role in criminal justice reform through reliance on data-driven interventions to reduce unnecessary reliance on incarceration while ensuring public safety. *See, e.g.*, *The Front End of the Criminal Justice System*, LAURA AND JOHN ARNOLD FOUND., http://www.arnoldfoundation.org/initiative/criminal-justice/crime-prevention/ (investing in "data, analytics and technology" to improve criminal justice decision making); *Criminal Justice*, MACARTHUR FOUND. (Oct. 2016), https://www.macfound.org/programs/criminal-justice/strategy/ (investing in data analytics research). As such, they are both likely to continue pursuing the development and use of predictive tools.

[53] Tools are sometimes referred to as "generations" because tool capabilities have evolved over time. *See* Melissa Hamilton, *Risk-Needs Assessment: Constitutional and Ethical Challenges*, 52 AM. CRIM. L. REV. 231, 236–39 (2015) (describing first- through fourth-generation risk tools). Generation delineation is not important to understanding risk assessment tools for this discussion. *See* Monahan & Skeem, *supra* note 7, at 499 ("In our view, distinctions between risk and needs (and associated generations of tools) create more confusion than understanding. Basically, tools differ in the sentencing goal they are meant to fulfill and in their emphasis on variable risk factors.").

[54] HARRIS ET AL., *supra* note 50, at 126 (in developing VRAG, the tool designers' goal was "an actuarial instrument to predict which offenders would commit at least one additional act of criminal violence given the opportunity").

[55] *See id.* at 137 (explaining impetus to develop the Sexual Offender Risk Appraisal Guide, which focuses on "the risk of violent recidivism among sex offenders," specifically); *Static-99/Static-99R*, STATIC99 CLEARINGHOUSE, http://wwww.static99.org (stating that "Static-99/R is the most widely used sex offender risk assessment instrument in the world").

[56] *See* EDWARD J. LATESSA ET AL., UNIV. OF CINCINNATI, THE OHIO RISK ASSESSMENT SYSTEM MISDEMEANOR ASSESSMENT TOOL (ORAS-MAT) AND MISDEMEANOR SCREENING TOOL (ORAS-MST) (2014), https://ext.dps.state.oh.us/OCCS/Pages/Public/Reports/ORAS%20MAT%20report%20%20occs%20version.pdf (predicting recidivism of misdemeanor offenders). A series of tools also predict an offender's tendency toward psychopathy and other dynamic characteristics like anger, which are outside the scope of this Article's focus. *See, e.g.*, ROBERT D. HARE, HARE PCL-R: HARE PSYCHOPATHY CHECKLIST-REVISED (2d ed. 2003) (describing creation of the psychopathy checklist); David J. Simourd, *The Criminal Sentiments Scale-Modified and Pride in Delinquency Scale: Psychometric Properties and Construct Validity of Two Measures of Criminal Attitudes*, 24 CRIM. JUST. & BEHAV. 52 (1997) (describing the link between criminal attitude and conduct).

factors to predict risk of re-offense only.[57] Other tools seek both to estimate the offender's risk of recidivism and to provide insight about interventions that could reduce risk.[58] Such tools, like the LSI-R, include static and dynamic risk variables, making them useful for sentencing and correctional use.[59]

The following sections describe how entities construct actuarial risk tools, focusing on key decisions of significance to sentencing law and policy. Section A discusses how entities choose to predict recidivism risk. It examines creation of the data set, development of the statistical model underlying any actuarial risk tool, selection of the predictive factors, and creation of the model. Section B discusses how entities choose to produce risk tools for use in the criminal justice system. It examines how tool developers convey models through various mechanisms and how they choose to translate quantitative model outcomes into qualitative recidivism risk scores and categories.

A. Predicting Recidivism Risk

No predictive tool is better than the data set from which it originates.[60] Data collection choices, like where and how to collect data and how to assemble a data set, provide the foundation for actuarial tools developed to assess recidivism risk. These decisions have a significant effect on the outcomes of the tools. As Professor Kate Crawford explains, "Data and data sets are not objective; they are creations of human design."[61] Data sets, she notes, are "intricately linked to physical place and human culture."[62] In other words, data only tells the story of the places from where it derives. To appreciate the scope

[57] Melissa Hamilton, Back to the Future: *The Influence of Criminal History on Risk Assessments*, 20 BERKELEY J. CRIM. L. 75, 92–93 (2015). For example, the VRAG uses twelve variables to assess recidivism risk. *Id.* at 93.

[58] *See* JAMES BONTA & D.A. ANDREWS, THE PSYCHOLOGY OF CRIMINAL CONDUCT 67 (6th ed. 2017).

[59] Hamilton, *supra* note 57, at 93–94. "Static" factors include those risk variables that cannot be changed, like age, gender, and criminal history. *See* D.A. Andrews & James Bonta, *Rehabilitating Criminal Justice Policy and Practice*, 16 PSYCHOL. PUB. POL'Y & L. 39, 45–46 (2010); *see also* Tonry, *supra* note 31, at 172 (noting that several static factors are actually variable markers, meaning that they are fixed at time of assessment, but subject to change). "Dynamic" factors include variables that are mutable in nature, like addiction and antisocial behavior. *See* Kelly Hannah-Moffat, *Actuarial Sentencing: An "Unsettled" Proposition*, 30 JUST. Q. 270, 275 (2013).

[60] *See* Stephen D. Gottfredson & Laura J. Moriarty, *Statistical Risk Assessment: Old Problems and New Applications*, 52 CRIME & DELINQ. 178, 183 (2006).

[61] Kate Crawford, *The Hidden Biases in Big Data*, HARV. BUS. REV. (Apr. 1, 2013), https://hbr.org/2013/04/the-hidden-biases-in-big-data.

[62] *Id.*

and limitations of a risk assessment tool, one must begin by considering the contours of its foundational data set.

The following subsections examine the great discretion entities developing risk tools exercise in creating risk prediction models. First, developers select the data. Second, developers must define "recidivism" in terms of a measurable target variable, like arrest. Third, developers select the predictive factors to observe in a data set. These factors may originate from empirical literature on recidivism, but not necessarily. Fourth, developers construct the predictive model. Decisions about the data to collect, the recidivism event to observe, and the risk factors selected have great import to understanding what and how a resulting recidivism risk tool predicts.

1. Collecting the Data

To predict risk of recidivism, tool creators collect data on people charged or convicted of crimes in the past as a base population.[63] Developers obtain this data from a variety of sources. They may conduct a study to observe the behavior of the selected individuals themselves over time.[64] They may obtain information directly from government agencies for repurposing.[65] They may simply collect publicly available information for repurposing.[66] Developers decide how to select the base population for their data sets. This decision creates the world within which statistical models generate predictions.

---

[63] Although developers could collect information about any set of individuals, see *infra* note 83, they tend to collect information about individuals *charged* or *convicted* of a crime in the past. *See, e.g.,* EDWARD LATESSA ET AL., UNIV. OF CINCINNATI, CREATION AND VALIDATION OF THE OHIO RISK ASSESSMENT SYSTEM: FINAL REPORT 13–14 (2009), http://www.ocjs.ohio.gov/ORAS_FinalReport.pdf (collecting data based on "adult[s] charged with a criminal offense" for both the pretrial and postconviction risk assessment tools); VA. CODE ANN. § 17.1-803 (West 2013) (directing the Virginia Sentencing Commission to develop a risk assessment instrument for sentencing "based on a study of Virginia felons").

[64] *See, e.g.,* LATESSA ET AL., *supra* note 63, at 15–16.

[65] *See, e.g.,* PA. COMM'N ON SENTENCING, RISK/NEEDS ASSESSMENT PROJECT: INTERIM REPORT 2 ON RECIDIVISM STUDY: INITIAL RECIDIVISM INFORMATION 1–2 (2011), http://pcs.la.psu.edu/publications-and-research/research-and-evaluation-reports/risk-assessment/phase-i-reports/interim-report-2-recidivism-study-initial-recidivism-information [hereinafter INTERIM REPORT 2] (collecting arrest information from state police and date of release from prison or probation from the department of corrections).

[66] This information may be collected from a private vendor. *See* FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY 11–12 (2014), https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf (data brokers can collect data from state and local governments for repurposing); *see also* Andrew D. Selbst, *Disparate Impact in Big Data Policing,* 49 GA. L. REV. (forthcoming 2017) (discussing the risk of errors in data brokers databases).

Developers tend to derive the base population from one of two sources: individuals observed upon release from prison or those referred to probation services. The ORAS includes two tools used at sentencing—the Community Supervision Tool (ORAS-CST) and the Misdemeanor Assessment Tool (ORAS-MAT).[67] Dr. Edward Latessa and his team of researchers designed the ORAS-CST using 681 adults charged with any criminal offense (felony or misdemeanor) and referred to probation services between September 2006 and February 2007.[68] The Northpointe General Recidivism Risk tool used 30,000 presentence investigation reports and probation intake cases collected from prison, parole, jail and probation sites.[69] The Pennsylvania Commission on Sentencing's risk tool relied on information pulled from the presentencing investigation reports of 41,563 individuals collected and maintained by the state's sentencing commission between 2004 and 2006.[70] The VRAG used data from two studies on recidivism conducted in Canada.[71]

Underlying data originate from a variety of locations as well. Commercial tools tend to derive their data from selected offenders in narrowly defined regions. For example, the multi-state, commercial tools developed by Northpointe, Inc. generated a data set from prisoners in several unidentified correctional facilities in the northeastern region of the United States.[72] State-specific risk assessment tools introduce geographic limits to their base population as well. For example, Pennsylvania's Commission on Sentencing conducted a recidivism study to provide the basis for the development of the state-specific sentencing tool.[73] The Commission's data set is composed of offenders from just four of the sixty-seven counties in the state—Centre,

---

[67] *See* Ohio Judicial Conference Cmty. Corr. Comm., *Policy Statement on the Ohio Risk Assessment System and Risk and Needs Assessment Tools*, OHIO JUD. CONF. 1 (Mar. 20, 2015), http://ohiojudges.org/Document.ashx?DocGuid=9e4c2814-6ffa-4018-9156-88fea13bf95e.

[68] LATESSA ET AL., *supra* note 63, at 14. Latessa and his team designed the ORAS-MAT in 2014 using a subset of the data pulled for creation of the ORAS-CST. *See* LATESSA ET AL., *supra* note 56, at 8.

[69] *See* NORTHPOINTE, INC., *supra* note 45, at 11; *see also COMPAS Risk & Need Assessment System: Selected Questions Posed by Inquiring Agencies*, NORTHPOINTE, INC. (2012), http://www.northpointeinc.com/files/downloads/FAQ_Document.pdf (providing an overview of the many norm groups available).

[70] INTERIM REPORT 2, *supra* note 65, at 1–2.

[71] *See* HARRIS ET AL., *supra* note 50, at 125.

[72] NORTHPOINTE, INC., PRACTITIONER'S GUIDE TO COMPAS 15 (2012), http://www.northpointeinc.com/files/technical_documents/FieldGuide2_081412.pdf.

[73] PA. COMM'N ON SENTENCING, RISK/NEEDS ASSESSMENT PROJECT: INTERIM REPORT 1: REVIEW OF FACTORS USED IN RISK ASSESSMENT INSTRUMENTS 1 (2011), http://pcs.la.psu.edu/publications-and-research/research-and-evaluation-reports/risk-assessment/phase-i-reports/interim-report-1-review-of-factors-used-in-risk-assessment-instruments [hereinafter INTERIM REPORT 1].

Berks, Philadelphia, and Delaware.[74] By comparison, the Ohio study included offenders from fourteen of the eighty-eight counties across the state.[75] There, developers specifically sought to create geographic diversity in the data set.[76]

2. Defining Recidivism

Selecting the base population for observation is only part of the initial data collection process. To glean information from that base population, developers must specify the outcomes they wish to study and the key variables they wish to observe.[77] This requires developers to translate a problem—here, recidivism—into a formal question about variables.[78] Framing this question requires that developers understand the objectives and requirements of the problem and convert this knowledge into a data problem definition.[79] It is a "necessarily subjective process," requiring developers to finesse a social dilemma such that a computer can automate a responsive answer.[80]

Developers frame this question around what they would like to know at sentencing: whether this person will commit a crime in the future. They translate the problem of public safety into a series of questions about the reoccurrence of criminal behavior and timing.[81] For instance, what events constitute "recidivism"? How far into the future should a tool predict this occurrence?[82] Developers resolve these issues by creating a simple yes–no question for observation in the data set.

Yet defining recidivism is less intuitive and more subjective than it may appear. Recidivism means the reoccurrence of criminal behavior by an individual.[83] This is not necessarily a binary outcome—many events could

---

[74] *Id.* at 8.

[75] LATESSA ET AL., *supra* note 63, at 13; *see Ohio County Map*, MAPS OF WORLD, http://www.mapsofworld.com/usa/states/ohio/ohio-county-map.html (displaying the eighty-eight counties in Ohio) (last updated Aug. 25, 2017).

[76] *See* LATESSA ET AL., *supra* note 63, at 12.

[77] *See* Barocas & Selbst, *supra* note 14, at 678.

[78] *See id.*

[79] *Id.*

[80] *Id.*

[81] Joan Petersilia, *Recidivism, in* ENCYCLOPEDIA OF AMERICAN PRISONS 382 (McShane & Williams eds., 1996).

[82] *Id.*; *see also* Robert Weisberg, *Meanings and Measures of Recidivism*, 87 S. CAL. L. REV. 785, 787–88 (2014) (discussing why we care about recidivism).

[83] Petersilia, *supra* note 81, at 382. Recidivism need not be limited to individuals previously convicted of a crime. However, as Dr. Joan Petersilia notes, "It is much easier to observe . . . [recidivism] among known offenders" compared to the population at large. *Id.* It is also an important goal of the criminal justice system

demonstrate recidivism. Regarding those already entangled in the criminal justice system, these events can vary in public safety significance. A fully adjudicated event may be the trigger. For example, someone may be considered a recidivist when *convicted of a new crime* for a particular type of offense, such as a violent offense, a property offense, or any offense whatsoever.[84] A criminal justice event that is not fully adjudicated may be the trigger as well. For example, someone may be considered a recidivist if *arrested and formally charged* for a particular offense.[85] The trigger may be merely *any new arrest regardless whether formally charged* for a particular type of offense.[86] For those individuals already under criminal justice supervision, the triggering event may be the *revocation of probation* leading to return to jail or prison for a new offense.[87] This may or may not include less serious occurrences such as a technical violation of the terms of parole or probation for administrative reasons, like failure to meet with a probation officer on time, failure of a drug test, or failure to pay criminal justice debt.[88]

Most tools rely on arrest as the measure of recidivism, although some variation exists within this principle across tools. For example, VRAG uses "any new criminal charge for a violent offense."[89] The explanation behind VRAG's definition provides helpful insight as to the types of choices developers face in defining the recidivism event. Researchers developing VRAG identified these crimes of violence using Canadian legislation combined with a bit of their own judgment calls, too.[90] For example, the tool includes all instances of assault whether physical contact occurred or not.[91] It also uses all instances of sexual assault even if the state would classify such offenses as "nonviolent child molestation."[92] Additionally, as the base population may have been institutionalized during the study, the developers included subsequent violent acts that, "in the judgment of research assistants, would have resulted in criminal charges had the incident occurred outside an

---

more broadly to reduce recidivism among those who have been punished by the system previously. *Id.*; *see also* Eaglin, *supra* note 37, at 608–09 (discussing the increasing importance of recidivism rates in sentencing reform policy).

[84] Petersilia, *supra* note 81, at 384.
[85] *Id.*
[86] *Id.*
[87] *Id.*
[88] *See id.*; *see also* Eaglin, *supra* note 37, at 610 n.98.
[89] HARRIS ET AL., *supra* note 50, at 122.
[90] *See id.* at 122–23.
[91] *Id.* at 122.
[92] *Id.*

institution."[93] Any of these events would count as "violent recidivism" for purposes of the VRAG tool design.

More generalized tools use any arrest to indicate recidivism. In ORAS, researchers rely on "arrest for a new crime."[94] Northpointe's COMPAS tool uses any arrest, whether it is for a felony or a misdemeanor.[95] The Pennsylvania Commission's tool defines recidivism as re-arrest and re-incarceration on a technical violation for offenders sentenced to prison.[96]

Developers also determine how far into the future a risk assessment tool should predict when they assemble a data set.[97] The length of the underlying study constricts the amount of time into the future a tool may predict with any accuracy. Currently available tools vary in the length of time they purport to predict. The VRAG predicts recidivism over five years.[98] Developers used data collected over ten years to develop this tool.[99] Such depth of research is unique, as most tools predict recidivism over a shorter period of time. The ORAS tools estimate recidivism one year from the date of assessment.[100] Pennsylvania's Commission on Sentencing chose to track offenders for a three-year period during the data-collection phase of its recidivism study.[101] Based on its underlying data set, the COMPAS tool estimates likelihood of recidivism for approximately two years into the future.[102]

Decisions about what constitutes recidivism and the amount of time over which to collect data on its occurrence presents fundamental questions for tool design. Some tool creators offer explanations for their choices. Take ORAS, for example, where Dr. Edward Latessa concedes that he and his team collected data on a variety of potential outcome measures—including conviction, probation violation, and institutional rule infraction—before

---

[93] *Id.* at 123.

[94] LATESSA ET AL., *supra* note 63, at 15–16.

[95] *COMPAS Risk & Need Assessment System: Selected Questions Posed by Inquiring Agencies, supra* note 69.

[96] INTERIM REPORT 2, *supra* note 65, at 1.

[97] Petersilia, *supra* note 81.

[98] HARRIS ET AL., *supra* note 50, at 132.

[99] *Id.* at 131.

[100] LATESSA ET AL., *supra* note 63, at 16.

[101] INTERIM REPORT 2, *supra* note 65, at 1–2.

[102] See *COMPAS Risk & Need Assessment System: Selected Questions Posed by Inquiring Agencies, supra* note 69(noting the two-year limit); NORTHPOINTE, INC., *supra* note 45, at 11 (noting that underlying study was conducted between January 2004 and November 2005).

settling on arrest for a new crime.[103] He uses arrest largely because gathering information later in the criminal justice process would require a longer study period.[104] Additionally, he notes that arrest relates to public safety threats to the community.[105] Similarly, tool creators designing the VRAG explain the decision to use criminal charges instead of criminal convictions, noting that charges entail less measurement of error compared to convictions.[106] On the other hand, some tool creators do not explain their choices. The Pennsylvania Commission on Sentencing does not indicate why it chooses to use re-arrest as a metric for recidivism, nor does Northpointe, Inc.[107]

How developers choose to define recidivism touches on key sentencing policy decisions often left undecided by state actors. Whether someone in the underlying data set gets arrested during observation determines if recidivism occurred.[108] This information will be included whether or not the offense was officially exonerated, prosecutors declined to bring a charge, or a conviction was overturned.[109] Criminal justice scholars and actors have long debated the role of acquittals and uncharged behavior in sentencing.[110] Although constitutionally permissible to consider at sentencing,[111] states use acquittals and non-convictions to varying degrees.[112]

3. Selecting Predictive Factors

Tool developers identify potential factors that likely predict recidivism and then construct a statistical model relying on some of those factors. Once they run the model, the developers can determine which of these factors have a

---

[103] LATESSA ET AL., *supra* note 63, at 15.

[104] *Id.* at 15–16.

[105] *Id.* at 16. This is a less compelling point. Latessa and his team recognize this in the report, as they note that factors predictive of rule violation are also of concern to criminal justice personnel. *Id.* However, arrest helps "identify criminogenic needs that are likely to result in danger to the community." *Id.*

[106] *See* HARRIS ET AL., *supra* note 50, at 123.

[107] *See* NORTHPOINTE, INC., *supra* note 45; INTERIM REPORT 2, *supra* note 65.

[108] *See supra* notes 83–96.

[109] *See* Hamilton, *supra* note 57, at 104.

[110] *See* Talia Fisher, *Conviction Without Conviction*, 96 MINN. L. REV. 833 (2012) (challenging the binary portrayal of guilty versus not guilty). *Compare, e.g.*, Stephen Breyer, *The Federal Sentencing Guidelines and the Key Compromises upon Which They Rest*, 17 HOFSTRA L. REV. 1, 8–12 (1988) (explaining the U.S. Sentencing Commission's decision to design federal sentencing guidelines that rely on unadjudicated conduct), *with* Kevin R. Reitz, *Sentencing Facts: Travesties of Real-Offense Sentencing*, 45 STAN. L. REV. 523 (1993) (challenging policy reasons for relying on unadjudicated conduct at sentencing).

[111] *See* United States v. Watts, 519 U.S. 148, 154–55 (1997); Dowling v. United States, 493 U.S. 342, 354 (1990).

[112] *See* Reitz, *supra* note 110; *see also infra* Section III.A.

2017]
CONSTRUCTING RECIDIVISM RISK

statistically significant correlation with the event in interest. In other words, they design a predictive model to answer the question: when recidivism occurs, what other factors tend to be present? To answer that question, developers must decide which factors to observe in the data set and whether to fold any or all of those factors into the final model.[113]

In theory, empirical research on recidivism serves as the starting point for researchers looking to select predictive factors for testing. Criminologists began studying factors that predict recidivism with the development of parole prediction tools in the 1920s.[114] Since the decline of the Rehabilitative Ideal in the 1970s,[115] several criminologists have developed a robust body of research on how to reduce offender recidivism.[116] Through this research, these criminologists also identified a number of variables that appear to be robust predictors of recidivism.[117] For example, Dr. Paul Gendreau, Tracy Little, and Claire Goggin conducted a meta-analysis in 1996 to identify the most predictive recidivism risk factors.[118] They identified seventeen factors with statistically high correlations to recidivism, including criminal companions, antisocial behavior, criminogenic needs, adult criminal history, race, family rearing practices, social achievement, current age, substance abuse, family structure, intellectual functioning, family criminality, gender, and socio-economic status.[119] Such studies serve as the basis of the recent shift toward "evidence-based" practice and policy in correctional and sentencing reforms more broadly.[120]

Developers make judgment calls about what factors to study in a data set. Although empirical study may guide that decision, it is not required, and tools vary in their reliance on the literature. For example, COMPAS uses "key risk

---

[113] *See* Barocas & Selbst, *supra* note 14, at 688.

[114] HARCOURT, *supra* note 1, at 47.

[115] *See* Eaglin, *supra* note 31, at 222; Klingele, *supra* note 1, at 542–43.

[116] *See, e.g.*, D.A. ANDREWS & JAMES BONTA, THE PSYCHOLOGY OF CRIMINAL CONDUCT (5th ed. 2010); Francis T. Cullen & Paul Gendreau, *Assessing Correctional Rehabilitation: Policy, Practice, and Prospects*, CRIM. JUST. 2000, July 2000, at 109; Francis T. Cullen & Paul Gendreau, *From Nothing Works to What Works: Changing Professional Ideology in the 21st Century*, 81 PRISON J. 313 (2001).

[117] *See, e.g.*, Paul Gendreau et al., *A Meta-Analysis of the Predictors of Adult Offender Recidivism: What Works!*, 34 CRIMINOLOGY 575, 576 (1996); Don A. Andrews, *Recidivism Is Predictable and Can Be Influenced: Using Risk Assessments to Reduce Recidivism*, CORRECTIONAL SERV. CAN. (Mar. 5, 2015), http://www.csc-scc.gc.ca/research/forum/special/espe_a-eng.shtml.

[118] Gendreau et al., *supra* note 117; *see also* Oleson, *supra* note 11, at 1350.

[119] Gendreau et al., *supra* note 117, at 582–83.

[120] Hannah-Moffat, *supra* note 59, at 271; Klingele, *supra* note 1, at 556.

factors that have emerged from the recent criminological literature."[121] Some of the original risk prediction tools are developed by these social scientists and based on the accumulation of their data on the topic.[122] Developers designing Pennsylvania's predictive tool, however, did not visit the empirical research directly. Rather, they searched for factors prevalent in already existing risk tools then used whatever factors were readily available in their current data set.[123]

Developers likely "clean" the data as well, often introducing their assumptions into the data collection process. Because data sets originate from a variety of sources, information provided may be incorrect.[124] For example, a presentence investigatory report may state that a defendant is twenty-seven and recently completed a 100-year sentence for armed robbery. Researchers seeking to use that information will either "fix" the information or throw the defendant out of the data set. "Fixing" the information requires subjective judgments about what the information likely means. For example, that a twenty-seven-year-old person would have served a 100-year sentence is impossible, and so it is obviously incorrect. However, a researcher may assume that a defendant served ten years in prison for armed robbery. Another may assume that the defendant served one year in prison. Either assumption could be correct. Other information may help indicate which of these assumptions is more likely accurate, but there is no way to be certain whether those decisions were correct. Faced with the choice to exclude or fix the data, different developers may choose alternative responses. The significance of this defendant in the data set may be minimal, but one cannot be certain how many such judgment calls are made without detailed disclosure from the developers.

Having compiled a data set and selected potential predictive factors, developers create a statistical model to identify which of the potential factors have a statistically significant correlation with recidivism. That statistical model provides the basis for the resulting actuarial risk assessment tool used at sentencing, but a number of subjective design choices are embedded in the next stage of the process discussed below.

---

[121] *See* NORTHPOINTE, INC., *supra* note 45, at 2.

[122] *See supra* Section II.A (discussing the fact that Harris, Rice, and Quinsey created the VRAG and Andrews and Bonta created the LSI-R).

[123] *See* INTERIM REPORT 1, *supra* note 73, at 3–5.

[124] *See* Barocas & Selbst, *supra* note 14, at 684–85 (explaining how datasets may rely on incorrect or partial information).

4. Constructing the Model

Developers decide how many predictive factors to include in the final risk assessment tool. The VRAG uses twelve factors, including marital status, age, elementary school maladjustment, and living with biological parents to age sixteen.[125] COMPAS uses fifteen factors, including financial problems, vocation/educational background, family criminality, residential instability, and leisure.[126] The Virginia Sentencing Commission Risk tool uses eleven factors, including marital status, age, gender, employment status, and a series of criminal history related factors.[127]

Tool creators may use different statistical methods to "weigh" the variables relative to one another. Two statistical methods prevail. First, some researchers choose to use the Burgess Method. Conceived by Ernest Burgess in 1928, this statistical method gives all predictive variables in a model equal weight.[128] One using a predictive model that applies this method need only add each predictive factor together to determine the probability of recidivism. So, for example, if males are more likely to recidivate, then males get one point while females receive zero points for this variable.[129] A number of tools follow this model, including Pennsylvania's risk tool, wherein presence of a variable leads to the addition of a point and absence of a variable will be assigned a score of zero.[130]

As an alternative, some developers use the Weighted Burgess Method. Conceived by Professor Sheldon Glueck and research assistant Eleanor Glueck in 1930, this statistical method assigns different weights to different variables in the predictive model.[131] The weight assigned to certain variables depends on how predictive that variable may be.[132] For example, if researchers determine that those who have criminal records are more likely to recidivate in the predictive model, then the presence of prior convictions not only adds points to a score, but it adds more points than, say, gender, if the researchers find that

---

[125] Hamilton, *supra* note 3, at 14, 15 tbl.1.

[126] Oleson, *supra* note 11, at app. at 1400.

[127] *Id.* at app. at 1402.

[128] HARCOURT, *supra* note 1, at 51, 59.

[129] PA. COMM'N ON SENTENCING, RISK/NEEDS ASSESSMENT PROJECT: INTERIM REPORT 4: DEVELOPMENT OF RISK ASSESSMENT SCALE 3 (2012), http://pcs.la.psu.edu/publications-and-research/research-and-evaluation-reports/risk-assessment/phase-i-reports/interim-report-4-development-of-risk-assessment-scale/view [hereinafter INTERIM REPORT 4].

[130] *Id.* at 8.

[131] HARCOURT, *supra* note 1, at 60–61.

[132] *Id.*

variable is also influential, but less so.[133] Most tools, including, for example, those developed by the UCCI,[134] use the Burgess Method because it is easier for laypersons to apply and understand.[135] Tool developers now agree that the alternative prediction methods perform equally well.[136]

Most tool creators choose to place heavier weight on certain types of factors beyond the correlations identified in the statistical model to produce tools with similar accuracy and easier use at sentencing. This design practice is most prevalent in relation to criminal history factors used in risk assessment tools.[137] For example, some tools include the same criminal history event as multiple variables in a predictive model. The LSI-R, for example, measures "prior adult convictions, arrests, charges, parole violations, and other official records of violence."[138] The Pennsylvania risk model separately counts the total number of prior arrests and prior convictions, including those for the same offense.[139] In Minnesota's sexual recidivism tool, six of the nine categories of risk factors relate to criminal history, even when potentially derived from a single event.[140] In California's Static Risk Assessment, the tool counts eighteen separate criminal history factors, some of which overlap.[141] The criminal history factors may include a number of past convictions, past incarceration

---

[133] *See, e.g.*, INTERIM REPORT 4, *supra* note 129, at 3. The Pennsylvania Commission on Sentencing identifies a third statistical method—the predictive attribute analysis. This method centers on the most predictive factor for certain types of defendants (male versus female, for example). *Id.* Data researchers then assign weight to other predictive variables by predictive ability for that specific type of defendant. *Id.* This method is a more advanced version of the Weighted Burgess Method. Currently, some juvenile recidivism tools use this model. *See* DON M. GOTTFREDSON & HOWARD N. SNYDER, NAT'L CTR. FOR JUVENILE JUSTICE, THE MATHEMATICS OF RISK CLASSIFICATION: CHANGING DATA INTO VALID INSTRUMENTS FOR JUVENILE COURTS 12 (2005), https://www.ncjrs.gov/pdffiles1/ojjdp/209158.pdf.

[134] *See* LATESSA ET AL., *supra* note 63, at 17.

[135] INTERIM REPORT 4, *supra* note 129, at 8 (selecting the Burgess Method because it "was the most straightforward"). "[T]he central battle lines [in developing risk tools] were between the Burgess unweighted, multiple-factor model and the Glueck weighted, few-factor model." HARCOURT, *supra* note 1, at 68. *But see* INTERIM REPORT 4, *supra* note 129, at 3 (setting forth a third option: predictive attribute analysis).

[136] *See* INTERIM REPORT 4, *supra* note 129, at 8.

[137] *See* HARCOURT, *supra* note 1, at 72 (explaining the focus on criminal history factors); Harcourt, *supra* note 11, at 239 (explaining most risk tools converge on criminal history factors).

[138] *See* Hamilton, *supra* note 57, at 98 (citing N.S.W. DEP'T OF CORRECTIVE SERVS., LSI-R TRAINING MANUAL 13–15 (2002)).

[139] *See* PA. COMM'N ON SENTENCING, RISK/NEEDS ASSESSMENT: INTERIM REPORT 3: FACTORS THAT PREDICT RECIDIVISM FOR VARIOUS TYPES OF OFFENDERS 12 (2011), http://pcs.la.psu.edu/publications-and-research/research-and-evaluation-reports/risk-assessment/phase-i-reports/interim-report-3-factors-that-predict-recidivism-for-various-types-of-offenders/view [hereinafter INTERIM REPORT 3].

[140] *See* Hamilton, *supra* note 57, at 98 (citing Minnesota Sex Offender Screening Tool (2012)).

[141] *See id.*

sentences, violent or drug convictions,[142] and prior arrests.[143] Several tools include the number of alleged offenses, acquitted conduct, and juvenile deviance, even if the criminal justice system typically discounts such events.[144]

Tool creators decide which predictive factors observed in the statistical model will be included in the resulting actuarial risk tool. Generally, risk assessment tools include four[145] different categories of predictive factors: criminal history,[146] anti-social attitude,[147] demographics,[148] and socioeconomic status.[149] Tool creators tend to include predictive factors without reference to whether their use is regulated in state sentencing systems.[150] Certain factors observed in the underlying data set will be excluded because they bear little relation to recidivism.[151] Others, like race and gender, may have high statistical correlation, but due to constitutional and ethical concerns, will be excluded from the resulting actuarial tool.[152] Some factors may be highly predictive, but

---

[142] Oleson, *supra* note 11, at app.

[143] The Pennsylvania Sentencing Commission's draft risk assessment tool depends heavily on arrests, unlike most other tools developed. *See* Barry-Jester et al., *supra* note 3. The eight-factor risk tool predicts re-arrest, not reconviction, and almost 40% of the score's outcome depends on history of arrest including prior adult arrests, prior property arrests, and prior drug arrests. *See* PA. COMM'N ON SENTENCING, RISK/NEEDS ASSESSMENT PROJECT: INTERIM REPORT 8: COMMUNICATING RISK AT SENTENCING 7 (2014), http://pcs.la.psu.edu/publications-and-research/research-and-evaluation-reports/risk-assessment/phase-i-reports/interim-report-8-communicating-risk-at-sentencing/view [hereinafter INTERIM REPORT 8].

[144] *See* Hamilton, *supra* note 57, at 95–96.

[145] Starr, *supra* note 1, at 811 (noting that most tools include criminal-history variables, demographic variables, and socioeconomic variables). These are consistent with the core variables associated with criminogenic needs. Some refer to the "big six": antisocial values, criminal peers, low self-control, dysfunctional family ties, substance abuse, and criminal personality. Others refer to the "big four" variables: antisocial associates, attitudes, personality, and criminal history. Still others refer to the "central eight" variables: antisocial associates, attitudes, personality, criminal history, family/marital circumstances, school/work difficulties, antisocial leisure/recreation, and substance abuse. ANDREWS & BONTA, *supra* note 116, at 65–66; Oleson, *supra* note 11, at 1349 n.133.

[146] Criminal history is the most common recidivism risk factor. For a discussion of the focus on this factor over time, see HARCOURT, *supra* note 1, 56–72; see also Oleson, *supra* note 11, at 1355–56 (discussing the prevalence of adult criminal history amongst risk prediction tools).

[147] This may include a number of variables outside the defendant's control, including family relations, addictions, and mental conditions. Oleson, *supra* note 11, at 1362–64.

[148] Demographic variables include age, gender, and marital status. Starr, *supra* note 1, at 811.

[149] Socioeconomic factors include, for example, employment status, financial condition, residential stability, and living in neighborhoods with high crime. Oleson, *supra* note 11, at 1360–61.

[150] This author's research did not find a single reference to state sentencing decisions about which factors should be considered at sentencing in the discussion of actuarial risk tool development. Although lack of cross reference cannot be certain given limited transparency in tool creation, the ubiquitous silence on the topic is significant.

[151] *See* Oleson, *supra* note 11, at 1350–52.

[152] *See* Eaglin, *supra* note 31, at 216–17 (discussing the decline in using race in recidivism risk tools and the reasoning behind this trend); Oleson, *supra* note 11, at 1380–82 (stating that race is highly predictive). On

tool creators exclude them for ease of use.[153] Just because a predictive factor is observed in the statistical model and found to have a statistically significant correlation to recidivism does not mean it will be included in the resulting risk tool.[154] Rather, the statistical correlation that such factors have with recidivism provide only one of many factors that a tool creator may consider when choosing which predictive factors should be included in a tool.

Determining which factors should be considered at sentencing is a notoriously difficult policy choice that has divided scholars, judges, and lawmakers for decades. For example, the U.S. Sentencing Commission notably restricted the factors that can be considered relevant at sentencing.[155] As Professor John Monahan noted in 2006, "[w]ith the single exception of criminal history . . . virtually all of the variables that potentially could be used as scientifically valid risk factors for violence . . . are explicitly excluded from consideration in federal sentencing procedures."[156] This reality has changed slightly as the federal sentencing guidelines are now advisory,[157] and the Sentencing Commission has reconsidered its hardline exclusion of several factors pertinent to the defendant's personal background in recent years.[158] Nevertheless, this statement reflects the reality that most factors included in recidivism risk tools have been highly regulated in the sentencing context. This principle applies in the states as well as the federal system, where some state sentencing provisions similarly endeavor to regulate the factors that are considered under specific sentencing guidelines.[159] Although use of criminal history may be ubiquitous, the ability to consider other factors varies greatly amongst the states due to legislatively imposed guidelines or guidelines created by sentencing commissions.

---

the other hand, gender is frequently used in risk tools. For discussion of the problematic implications of including gender as a predictive factor in actuarial risk tools, see, for example, Starr, *supra* note 1, at 823–29.

[153] Oleson, *supra* note 11, at 1348–49.

[154] *Id.*

[155] *See* U.S. SENTENCING GUIDELINES MANUAL §§ 5H1.1–1.6 (U.S. SENTENCING COMM'N 2004); KATE STITH & JOSÉ A. CABRANES, FEAR OF JUDGING: SENTENCING GUIDELINES IN THE FEDERAL COURTS 74–75 (1998).

[156] Monahan, *supra* note 5, at 397–98.

[157] *See* United States v. Booker, 543 U.S. 220, 245 (2005).

[158] *See* U.S. SENTENCING GUIDELINES MANUAL app. C, vol. 3 (U.S. SENTENCING COMM'N 2011) (revising the guidelines to permit consideration of age, mental and emotional condition, physical condition or appearance, and military service).

[159] *See* Monahan, *supra* note 5, at 398–99 (discussing state and federal limitations). *See generally* DAN MARKEL ET AL., PRIVILEGE OR PUNISH: CRIMINAL JUSTICE AND THE CHALLENGE OF FAMILY TIES 15–16 (2009) (discussing various state approaches to consideration of family ties at sentencing, including limitations).

B. Creating a Risk Assessment Tool

After the design decisions are made, researchers create the actual risk assessment tool. The risk assessment tool applies the undisclosed algorithm that predicts recidivism consistently across cases. The algorithm reflects the normative judgments discussed above about what factors should count and how much when predicting recidivism. The algorithm produces an instantaneous quantitative outcome based on the information selected from the predictive model. This outcome suggests the numerical probability that the tool-defined recidivism event will occur with individuals sharing those same characteristics.[160] In other words, it ranks defendants according to likelihood of engaging in criminal behavior based on the behavior of the individuals in the underlying data set. Most risk tools translate that quantitative outcome into a qualitative "risk score" used by criminal justice actors at sentencing.

It is worth noting a few points about how criminal justice actors administer the tools to estimate a specific defendant's risk score. Developers create instructions for criminal justice actors to use when administering the tool.[161] Several tools require that probation officers or other intake personnel collect information via a structured interview with the defendant.[162] For example, the Indiana Risk Assessment System's Community Supervision Intake Assessment provides a five-page, structured interview questionnaire.[163] Some tools require that the defendant provide information voluntarily.[164] As an example, the COMPAS system permits the collection of offender information through official records data and a defendant's self-reporting.[165] Still others require no formal collection of information at all; instead, criminal justice administrators rely on publicly accessible data. As an example, the Laura and John Arnold

---

[160] *See* Hamilton, *supra* note 3.

[161] *See, e.g., Risk Assessment, supra* note 51; Judicial Conference of Indiana, *Policy for User Certification for the Indiana Youth Assessment System & Indiana Risk Assessment System*, IND. JUD. BRANCH (Aug. 25, 2011), http://www.in.gov/judiciary/cadp/files/prob-risk-iyas-iras-user-certification-2011.pdf.

[162] HARRIS ET AL., *supra* note 50, at 152; LATESSA ET AL., *supra* note 63, at 11–12. COMPAS offers an option—the defendant may fill out a self-report or the criminal justice administrator may conduct an interview. *COMPAS Risk & Need Assessment System: Selected Questions Posed by Inquiring Agencies, supra* note 69. Future tools may not require an interview at all. For example, the Laura and John Arnold Foundation is developing risk prediction tools that do not require a structured interview. *See Developing a National Model for Pretrial Risk Assessment, supra* note 52.

[163] *See* UNIV. OF CINCINNATI, *supra* note 51, at 2-3–2-8.

[164] *See id.*

[165] Structured interviews are available, but not required. *COMPAS Risk & Need Assessment System: Selected Questions Posed by Inquiring Agencies, supra* note 69; *see also* CASEY ET AL., *supra* note 41, app. at A-25.

Foundation developed a risk assessment tool that eliminates the interview process entirely for criminal justice administrators.[166] This practice will likely expand to other tools as well.[167]

These examples demonstrate the variety in administrative expectations required amongst current recidivism risk tools. Administrators rely on that information, along with official records and other collateral sources to complete the assessment tool.[168] Defendants may complete self-reports to complement that information. A criminal justice administrator collects any requisite information and puts it into a computer system or calculates the risk score by hand on a specified worksheet.[169]

Most risk tools translate the quantitative tool score into a qualitative risk classification to facilitate use.[170] To do this, thresholds are set into a statistical model's outcomes that divide offenders into different pools, or categories. The typical division is some version of low-, medium-, and high-risk offenders.[171] "Cut-off points" refers to the numerical scores that serve to divide the line between these categories of offenders.

An example illustrates this process. An actuarial risk tool may indicate that a defendant has a 10% chance of being arrested for a nonviolent act within one year of conducting the assessment, assuming nothing changes in that time frame. But actuarial tools do not communicate information from the predictive model in this way. Rather, the tools classify defendants into risk categories— meaning the tools indicate that a numerical score signifies a certain risk level.

---

[166] *See* Lauryn Gouldin, *Defining Flight Risk*, 85 U. CHI. L. REV. (forthcoming 2017) (describing the shift away from interviews as the basis of risk assessment tools in the pretrial detention context).

[167] *See supra* note 162.

[168] *See* CASEY ET AL., *supra* note 41, app. at A-56.

[169] For example, the Virginia, Ohio, and Indiana risk tools may be calculated by hand. UNIV. OF CINCINNATI, *supra* note 51, at 2-9–2-41. The COMPAS tools require a computer.

[170] Tool creators translate the numbers to words for a variety of reasons. Lay people may find numerical probabilities "unnatural and awkward" and "aesthetic[ally] revulsi[ve]" compared to language. PHILIP E. TETLOCK & DAN GARDNER, SUPERFORECASTING: THE ART AND SCIENCE OF PREDICTION 56 (2015). Tool creators want to develop tools that bridge the divide between data and practicalities with ease. *See id.* They may even want to represent a certain amount of surety in their calculations. *See id.* Recidivism classifications are familiar to criminal justice actors, whether those terms are backed by statistical data or not. *See* Jurek v. Texas, 428 U.S. 262, 275 (1976) ("[P]rediction of future criminal conduct is an essential element in many of the decisions rendered throughout our criminal justice system."). Sensitive to varying concerns and perspectives, tool creators likely translate the numerical risk scores into risk categories for broader use. *See* TETLOCK & GARDNER, *supra*, at 56.

[171] For a real-world example of the risk categories applied to statistical model outcomes, see INTERIM REPORT 8, *supra* note 143, at 4.

The 10% chance of recidivism may mean a low, medium, or high risk of recidivism depending on where tool creators introduce the cut-off points between classification levels.

Translating tool outcomes into risk categories is a highly subjective, policy-oriented process. This decision requires some expertise not only in what the tool is predicting,[172] but also in how society interprets the numerical outcome's meaning.[173] In short, where developers place cut-off points reflects a normative judgment about how much likelihood of risk is acceptable in society without intervention.[174] Using the example from above, is a 10% likelihood of engaging in criminal behavior an acceptable level of risk such that additional supervision is unnecessary? What about 15%? When should a level of risk shift from acceptable—meaning low risk—to unacceptable—meaning high risk? Tools vary in this judgment.

* * *

The normative judgments described above have real significance for a defendant at sentencing. Take as a hypothetical, defendant X and defendant Y. Defendant X is male, twenty-five, and lives in an urban center with a high crime rate. He dropped out of high school to take care of his pregnant girlfriend, and now has a young child. He is black. Defendant Y is female, forty, and lives in a rural community with a low crime rate. She finished high school, but fails to maintain a steady job due to drug addiction. Although she had a persistent shoplifting problem in high school, the stores never reported her to the police out of respect for her family. She is white. Both defendant X and defendant Y commit a crime, say theft. Both are convicted. The normative judgments described above will affect whether and how both of these defendants' are categorized for risk, which in turn can affect the length and amount of supervision each faces.[175]

Section A illustrates that factors like education, housing, employment, criminal history, and family ties can affect the outcome of the risk assessment tool. Developers make decisions about which of those factors matter, why, and how much. Section B illustrates that whether the tool results place defendant X in a higher risk category than defendant Y depends on how developers

---

[172]	See Barocas & Selbst, *supra* note 14, at 678–79.

[173]	See TETLOCK & GARDNER, *supra* note 170, at 53.

[174]	See Mayson, *supra* note 1 (discussing various levels of risk considered "high" for pretrial risk assessment tools).

[175]	See *supra* note 7.

translate the tool's numerical outcomes into qualitative risk categories. Defendant X and defendant Y may both be considered to have a high or low risk of recidivism subject to placement of the cut-off points between risk categories, even if the model ranks defendant X as more likely than defendant Y to engage in criminal behavior in the future. Developers must decide how to separate the tool outcomes to create risk categories. Together, these sections illustrate that the seemingly objective and neutral representation of one defendant as high risk and another as low risk is a matter of tool construction.

Part I demonstrates two further points as well. First, constructing actuarial risk tools for sentencing requires that developers make a variety of choices with great consequence to sentencing law and policy. These choices reflect normative judgments about what counts at sentencing and why. These choices also reflect larger normative judgments about how much risk society tolerates. States vary in how they may decide these choices, although the tools developed to estimate risk largely do not.

The second point develops more subtly throughout; namely, that it is difficult to ascertain these policy decisions on the face of the tool. Even though previous efforts to estimate risk at sentencing—like guidelines and mandatory penalties—made normative judgments about how to sentence and why, those choices were apparent on the face of the mechanized tool. For example, a mandatory penalty for an offense is triggered by a particular fact, like a prior conviction. With actuarial risk tools, normative judgments are more difficult or even impossible to discern. The following Part identifies threats these tools present to sentencing law and policy.

II. CONSTRUCTING TOOLS WITHOUT GUIDANCE: THREATS TO SENTENCING LAW AND POLICY

Tools constructed to estimate recidivism risk reflect numerous normative choices. There is no such thing as a "value-free" tool.[176] One might argue that developers do not need guidance on how to determine policy judgments embedded in risk tools' construction. Underlying this logic is the assumption that because the tools are data-driven, the information produced is objective, neutral, and valuable to society. This Part contests that assertion, thereby

---

[176] *See* BERK, *supra* note 41, at 6.

making the case for accountability measures for tools designed to predict recidivism risk.

Sections A through C identify three types of normative judgments implicated in decisions about how to construct actuarial risk tools: judgments about accuracy, equality, and the purpose of punishment. Currently, developers can make these decisions during tool construction without guidance. Yet, as section D will explain, developers have unique and diverging interests that shape tool construction and may, at times, result in conflicting or contradictory decisions compared to a state's sentencing policy.

A. Differing Notions of Accuracy

In the criminal justice system more broadly, there are two types of accuracy. Accuracy may mean reducing the number of people who committed crimes from evading punishment. It may also mean reducing the number of innocent people who have not committed crimes from wrongfully experiencing punishment. These concepts drive at "the twofold aim [of criminal justice] . . . that guilt shall not escape or innocence suffer."[177] Each of these types of accuracy comes with a cost. Ensuring that the guilty do not escape punishment at times means increasing the chance that the innocent will suffer. Similarly, ensuring that the innocent do not suffer means, at times, increasing the chance of letting the guilty escape punishment. At a theoretical level, society prefers the latter. [178] Yet whether a risk prediction tool should identify fewer or more defendants as a recidivism risk—and how many more—is a normative judgment for which no heuristic can apply and empirical accuracy measures cannot resolve.

Actuarial risk assessment tools strive toward two separate types of accuracy as measured by empirical validity studies. These measures are distinct from the criminal justice aims. On the one hand, there is predictive accuracy, meaning that a tool predicts what it portends to predict better than

---

[177] United States v. Nixon, 418 U.S. 683, 709 (1974) (alteration in original) (quoting Berger v. United States, 295 U.S. 78, 88 (1935)).

[178] Blackstone's famous adage, "that it is better that ten guilty persons escape, than that one innocent suffer," indicates a preference. See 4 WILLIAM BLACKSTONE, COMMENTARIES *352. This statement reflects the preference that, all things being equal, the system should protect the innocent from wrongful punishment even at the expense of letting the guilty go free. With the mechanization of criminal justice, this simple preference is placed in doubt at a practical level. See Roth, Trial by Machine, supra note 33, at 1252–53, 1267–69 (describing criminal mechanizations' uneven desire for a particular kind of accuracy that prevents lenience and mercy).

chance.[179] For example, a risk assessment tool is accurate if it successfully differentiates between those who experienced the outcome of interest and those who did not more than 50% of the time.[180] On the other hand, there is classification accuracy. A tool achieves this type of accuracy if "the average predicted recidivism rate is relatively equal to the actual rate of recidivism."[181] In the context of recidivism risk tools, this means that the predicted outcome event occurs as frequently as anticipated by the tool. Using an example from Professor Melissa Hamilton, if a tool estimates that 10% of persons categorized as moderate risk will recidivate and the actual observed recidivism rate of the moderate group in a validity study is about 10%, then the tool classifies risk accurately.[182]

Validity studies assessing predictive and classification accuracy offer useful, but limited, information regarding the value of a risk assessment tool at sentencing. For example, a popular method to assess predictive accuracy measures the area under the curve, or the AUC value.[183] This measurement indicates how well a risk assessment tool discriminates between recidivists and non-recidivists relative to the occurrence of the event of interest in the base data set and uninhibited by the policy decisions inherent to classifying outcomes into risk categories.[184] Yet the AUC value provides limited

---

[179] See, e.g., Slobogin, supra note 34, at 292.

[180] See Hamilton, supra note 3, at 24–26 (referring to this type of accuracy measure as "discrimination"); see also Slobogin, supra note 34, at 292. To measure predictive accuracy, developers submit tools to validity studies such as measuring the area under the curve (AUC), discussed below. See infra notes 183–184. An AUC value of .50 means that the tool predicts equally as well as chance. Slobogin, supra note 34, at 292. Most tools used today have AUC values between .60 to .80. Id. at 293.

[181] See Hamilton, supra note 3, at 24 (describing this type of accuracy measure as "calibration").

[182] Id.

[183] See Hamilton, supra note 3, at 26. Many researchers felt the other predictive accuracy measurements underrepresented the accuracy of actuarial risk tools because it was constrained by the base rate in a data set. Gottfredson & Moriarty, supra note 60, at 186 ("The problem in using any of these [other current validity] measures . . . is that the tool's apparent usefulness is highly dependent on the base rate, [as well as] the selection ratio . . . ."); see also Paul R. Falzer, Valuing Structured Professional Judgment: Predictive Validity, Decision-making, and the Clinical-Actuarial Conflict, 31 BEHAV. SCI. & L. 40, 43–44 (2013); R. Karl Hanson & David Thornton, Improving Risk Assessments for Sex Offenders: A Comparison of Three Actuarial Scales, 24 LAW & HUM. BEHAV. 119, 125 (2000).

[184] As Professor Melissa Hamilton explains, "The correct interpretation of the AUC (for a recidivism risk tool) is 'the probability that a randomly selected individual who committed an [act of recidivism] . . . received a higher risk classification than a randomly selected individual who did not' reoffend." Hamilton, supra note 3, at 25 (citing Jay P. Singh et al., Measurement of Predictive Validity in Violence Risk Assessment Studies: A Second-Order Systematic Review, 31 BEHAV. SCI. & L. 55, 64 (2013)) (alteration in original). "The ROC area has advantages over other commonly used measures of predictive accuracy . . . because it is not constrained by base rates or selection ratios . . . ." Hanson & Thornton, supra note 183, at 125 (citation omitted). The AUC

information about the value of the tool's outcome for sentencing. For example, just because a tool has a high AUC value does not mean that a defendant with a high risk score will eventually engage in the recidivism event of interest.[185]

But the limitations go beyond what the predictive accuracy measurement says about the tool's results. This method of empirical validation says nothing about the construction of the tool itself. For example, predictive accuracy measures say nothing about the legitimacy of risk factors that developers choose to include or exclude from the tool as a matter of sentencing policy.[186] Did the tool use information that is prohibited in a state by law? Predictive accuracy measures do not provide an answer. Nor does this measure of tool accuracy say anything about the propriety of using a particular definition of recidivism for a particular use in the justice system. In other words, does a jurisdiction care whether a defendant is likely to be rearrested or reconvicted for particular behavior, and if so, for what? Validity studies assessing the predictive accuracy of a tool cannot answer these questions because they are inherently normative.

Similarly, validity studies provide little guidance regarding how developers divide a risk tool's outcomes into various risk categories. Recall that any risk assessment tool translating a predictive score into a qualitative risk assessment requires that tool developers separate results into risk categories or "bins" like high, medium, or low risk. Predictive accuracy measures like the AUC value provide no insight into whether the cut-off points located between high, medium, and low recidivism risk categories accurately calibrate with actual outcomes in the real world.[187] Yet even validity studies that measure a tool's classification accuracy cannot resolve the normative judgment about how many defendants *should* fall into each category for sentencing. Because the tools are designed to study human behavior, they inevitably result in errors, meaning instances when the predicted outcome does not occur.[188] While

---

value is a fraction obtained from the "ROC value," meaning receiver operating characteristic curve, referenced by Hanson and Thornton. *See* Hamilton, *supra* note 3, at 25.

[185] Hamilton, *supra* note 3, at 25.

[186] *See, e.g.*, R. Karl Hanson & Philip D. Howard, *Individual Confidence Intervals Do Not Inform Decision-Makers About the Accuracy of Risk Assessment Evaluations*, 34 LAW & HUM. BEHAV. 275, 281 (2010) ("[T]he judgment concerning the credibility of the risk assessment procedure . . . is fundamentally qualitative.").

[187] *See, e.g.*, Hamilton, *supra* note 3, at 27 ("An AUC [value] can be far above .50 even if the tool is not well-calibrated (e.g., the percentage of predicted outcomes is significantly different than the proportion of the actual outcomes).").

[188] *See, e.g., id.* at 25; Starr, *supra* note 1, at 843.

studies demonstrating classification accuracy consider actual outcomes within the risk bins, the information produced by the tools cannot resolve the tradeoffs inherent to the construction of those bins. Said differently, this information says nothing about whether and when classification errors impose unbearable costs to society when used at sentencing. Addressing this question requires consideration of normative judgments about the costs of miscategorizing defendants.

Classification accuracy measurements turn on the placement of cut-off points between risk categories. Implicit in this placement lies a normative judgment about the cost of errors.[189] At sentencing, the costs associated with error differ greatly. Failing to identify a defendant as high risk who goes on to engage in criminal behavior, for example theft, in the future creates a false negative. It may result in that defendant being released into the community without more intense criminal justice supervision. This failure imposes a cost to the victim and his family, to society, and to criminal justice actors. The victim and his family may experience the physical, emotional, and psychological harms from the crime. Society incurs the cost of a crime that could have been prevented. Criminal justice actors incur a cost to their credibility. On the other hand, identifying a person as high risk who would not commit a crime in the future—a false positive—also imposes costs. There, the risk classification may lead to additional and unnecessary criminal justice supervision. The defendant experiences an infringement to her freedom that can hinder her reintegration into society. Society incurs the economic costs of unnecessarily diverting limited criminal justice resources toward the defendant. It incurs the social costs as well, including the increased likelihood of a recidivism event. The criminal justice system also incurs a cost to its credibility. Although not an innocent person, the erroneously classified person experiences more punishment than she otherwise should.

No "neutral" answer exists to balance the costs described here,[190] yet addressing these costs relates to traditional criminal justice notions of accuracy. Punishment, like guilt, imposes costs that must be balanced in the justice system. A lower risk score cut-off point (for example, categorizing anything more than a 10% likelihood of engaging in criminal behavior in the

---

[189] *See* Richard Berk, *Balancing the Costs of Forecasting Errors in Parole Decisions*, 74 ALB. L. REV. 1071, 1074 (2011).

[190] For more on the existence of normative judgments in the construction of risk tools, see, for example, *id.*; Berk & Hyatt, *supra* note 41; Mayson, *supra* note 1.

future as high risk) captures more of what you want—true positives, meaning defendants who would recidivate—and more of what you do not want—false positives.[191] A higher cut-off point (for example, categorizing anything over a 90% likelihood as high risk) does the opposite. Where developers place the cut-off points between various likelihoods of recidivism will affect how many defendants a tool identifies as significant risk. That determination can subject defendants to more or less supervision in the criminal justice system regardless of whether they actually engage in criminal behavior in the future. Reasonable minds can and do differ regarding what is the right placement to balance the costs of error at sentencing.

Though the placement of cut-off points relates to both empirical validity studies and criminal justice accuracy, the law provides no assurance that this construction decision reflects the values of the jurisdictions where tools are adopted.[192] Indeed, lawmakers rarely offer guidance on how to construct actuarial risk tools at all. More often, legislators and policymakers advocating for risk-based sentencing simply point to validity studies from the forensic science field as a measure of tool quality.[193]

Of course, tool developers do not suggest that risk tools will eliminate errors; rather, they suggest that the tools can improve judicial decision making by reducing errors in determining a defendant's risk level.[194] But when a risk

---

[191] *See, e.g.*, Berk, *supra* note 189, at 1074–75 (discussing the relative costs of error in parole forecasting); Hamilton, *supra* note 3, at 33–35 (discussing costs of error at sentencing).

[192] While risk estimates produced by clinicians were subject to testing through scientific boards and examination on the stand, actuarial risk tools are not currently subjected to this rigorous testing.

[193] *See, e.g.*, CASEY ET AL., *supra* note 7, at 14–18 (urging local validation of risk assessment tools to ensure reliability); *see also* DAVID FARABEE ET AL., CAL. DEP'T OF CORR. & REHAB., COMPAS VALIDATION STUDY: FINAL REPORT 3–4 (2010), http://www.cdcr.ca.gov/Adult_Research_Branch/Research_Documents/ COMPAS_Final_Report_08-11-10.pdf (assessing California's general recidivism risk scale as acceptable); SHARON LANSING, DIV. OF CRIMINAL JUSTICE SERVS., NEW YORK STATE COMPAS-PROBATION RISK AND NEED ASSESSMENT STUDY: EXAMINING THE RECIDIVISM SCALE'S EFFECTIVENESS AND PREDICTIVE ACCURACY i (2012), http://www.criminaljustice.ny.gov/crimnet/ojsa/opca/compas_probation_report_2012.pdf (assessing validity of risk tool in New York); BRIAN J. OSTROM ET AL., NAT'L CTR. FOR STATE COURTS, OFFENDER RISK ASSESSMENT IN VIRGINIA: A THREE-STAGE EVALUATION: PROCESS OF SENTENCING REFORM, EMPIRICAL STUDY OF DIVERSION & RECIDIVISM, BENEFIT-COST ANALYSIS 8 (2002), http://www.vcsc.virginia.gov/risk_ off_rpt.pdf (endorsing Virginia's risk assessment tool on the basis of validity studies); JENNIFER L. SKEEM & JENNIFER ENO LOUDEN, CAL. DEP'T OF CORR. & REHAB., ASSESSMENT OF EVIDENCE ON THE QUALITY OF THE CORRECTIONAL OFFENDER MANAGEMENT PROFILING FOR ALTERNATIVE SANCTIONS (COMPAS) 28 (2007), http://www.cdcr.ca.gov/Adult_Research_Branch/Research_Documents/COMPAS_Skeem_EnoLouden_Dec_2 007.pdf (recommending California not use COMPAS without more evidence).

[194] *See* Jennifer L. Skeem & John Monahan, *Current Directions in Violence Risk Assessment*, 20 CURRENT DIRECTIONS PSYCHOL. SCI. 38 (2011).

assessment tool can impact a criminal justice outcome like sentencing, it is important to look beyond just validity studies to determine the value of the tool at sentencing. Validity studies cannot resolve the difficult normative judgments inherent to constructing risk tools used in pursuit of the twin aims of the criminal justice system. Only the communities where the tool will be applied can do that.

B. Compromising Equality

Actuarial risk tools may compromise equality as a matter of construction. "Equality" has various meanings.[195] This section engages with the concept as a matter of equal opportunity for a tool to estimate a defendant as low risk. Using race to demonstrate the point,[196] this section will illustrate that risk tools prevent this type of equality among defendants based on societal realities potentially overlooked during the tool-construction process.

Most risk tools estimate recidivism risk as the likelihood of engaging in behavior that will lead to arrest, not conviction. Arrest is an action taken by police officers under authority of the state.[197] These are some of the least procedurally protected instances of contact with the criminal justice system.[198]

---

[195] *See* Richard Berk et al., *Fairness in Criminal Justice Risk Assessments: The State of the Art*, CORNELL U. LIBR. 12–15 (May 30, 2017), https://arxiv.org/pdf/1703.09207.pdf (describing various meanings of accuracy and equality from a statistical perspective); Sandra G. Mayson, *Bias In, Bias Out: Criminal Justice Risk Assessment and the Myth of Race Neutrality* (unpublished manuscript) (on file with author) (describing various meanings of accuracy and equality from a legal perspective).

[196] The same analysis applies to ethnic disparities prevalent in the criminal justice system as well.

[197] "[T]he police arrest a suspect whenever they, on the basis of suspicion that he has committed a criminal offense or violation, (1) take him into custody by handcuffing or otherwise depriving him of his freedom; (2) transport him to a police station, jail, or detention facility; (3) process him by creating a permanent record of the arrest, taking identifying information, including photographs, fingerprints, and the like; and (4) detain him until either he is released or his arrest is subjected to judicial review." Rachel A. Harmon, *Why Arrest?*, 115 MICH. L. REV. 307, 311 (2016).

[198] A police officer may arrest someone based upon probable cause to believe that a person committed a crime. Tennessee v. Garner, 471 U.S. 1, 7 (1985). Probable cause provides a "relatively low threshold" for police intervention. *See* Rachel A. Harmon, *The Problem of Policing*, 110 MICH. L. REV. 761, 779 (2012) ("[P]robable cause ensures only that there is a reason to arrest the individual, not that the arrest is a necessary or effective means of enforcing the law or preventing disorder."); Eisha Jain, *Arrests as Regulation*, 67 STAN. L. REV. 809, 818 (2015). While less scrutinized criminal enforcement events also occur—like *Terry* stops or traffic stops—arrests result in criminal records that can follow a defendant for life. Harmon, *supra* note 197, at 312 ("Unlike many other encounters with the police, a suspect who is arrested and booked faces practical, reputational, and privacy consequences that persist whether or not he is subject to further legal proceedings."). *See generally* Jain, *supra*, at 820–25 (describing impact of arrest).

Of the twelve million people arrested every year,[199] many are not ultimately convicted. Explanations abound for this—prosecutors may choose not to pursue charges,[200] evidence may not support the charge, prosecutors may not secure convictions,[201] and, most simply, the defendant may be innocent.[202]

Arrests occur disproportionately against minorities, and in particular, against black men. [203] Black men come into contact with the criminal justice system more frequently[204] and from an earlier age.[205] But more frequent contact with the justice system does not necessarily mean higher risk to the public. Much of this contact comes from heightened scrutiny, not necessarily more criminal wrongdoing. Nowhere is this more apparent than in the context of drug crimes, where police disproportionately arrest blacks even though blacks use drugs at similar rates as other races.[206] The decision to use arrests as a predictive factor and event of interest thus lies in some other purpose than public safety.

Actuarial risk tools rely on a number of other factors beyond contact with the criminal justice system that also disproportionately disadvantage communities of color. Popular risk tools rely on factors like education, employment, parents with criminal history, and marital status—which may disadvantage black defendants given cultural realities and structural barriers in

---

[199] *Crime in the United States 2012*, UNIFORM CRIME REPORTING: FBI (2012), https://ucr.fbi.gov/crime-in-the-u.s/2012/crime-in-the-u.s.-2012/persons-arrested.

[200] *See* Stephanos Bibas & Richard A. Bierschbach, *Integrating Remorse and Apology into Criminal Procedure*, 114 YALE L.J. 85, 128 (2004) ("[P]rosecutors can choose whether to accept police officers' recommendations and pursue those charges.").

[201] *See* Jenny E. Carroll, *Nullification as Law*, 102 GEO. L.J. 579, 604–09 (2014) (explaining that juries may refuse to convict a defendant); Anna Roberts, *Dismissals as Justice*, ALA. L. REV. (forthcoming 2017) (showing that judges may dismiss prosecutions).

[202] Josh Bowers, *Legal Guilt, Normative Innocence, and the Equitable Decision Not to Prosecute*, 110 COLUM. L. REV. 1655, 1680–84 (2010).

[203] Blacks are arrested at higher rates than whites or Hispanics. *See* JESSICA EAGLIN & DANYELLE SOLOMON, BRENNAN CTR. FOR JUSTICE, REDUCING RACIAL AND ETHNIC DISPARITIES IN JAILS: RECOMMENDATIONS FOR LOCAL PRACTICE 17–18 (2015), https://www.brennancenter.org/sites/default/files/publications/Racial%20Disparities%20Report%20062515.pdf. Even disparities in convictions cannot explain the disparity in arrests. *Id.* at 18–19. This is particularly true in the context of drug crimes, where African Americans comprise 31% of those arrested for drug law violations despite making upon only 13% of the U.S. population and using drugs at similar rates as other races. DRUG POLICY ALL., THE DRUG WAR, MASS INCARCERATION AND RACE 1 (2016), http://www.drugpolicy.org/sites/default/files/DPA%20Fact%20Sheet_Drug%20War%20Mass%20Incarceration%20and%20Race_%28Feb.%202016%29_0.pdf.

[204] *See* DRUG POLICY ALL., *supra* note 203.

[205] *See* MICHAEL TONRY, MALIGN NEGLECT: RACE, CRIME, AND PUNISHMENT IN AMERICA 29–30 (1995).

[206] *See* DRUG POLICY ALL., *supra* note 203.

society.[207] Black and Latino applicants searching for a job without a criminal record fare no better than white applicants just released from prison. One in fifteen black children born today has an incarcerated parent, as compared to one in 111 white children.[208] Children of incarcerated parents tend to suffer from learning and behavioral problems at higher rates than peers whose parents are not incarcerated.[209] Even without the reliance on criminal history, other variables disproportionately affect racial minorities with such frequency that tools relying on those factors will classify minorities as higher risk.

These realities combine to demonstrate that risk tools by design will more frequently classify minorities as higher risk. Consider the ongoing debate regarding COMPAS risk tools as an example. ProPublica recently conducted a study on the use of commercially designed tools.[210] There, journalists found through statistical analysis that black defendants evaluated by COMPAS tools were more likely to be incorrectly labeled as higher risk without committing a future crime in the requisite time period, as compared to white defendants who were more likely to be incorrectly labeled as lower risk but actually committed crimes in the same time period.[211] This disparity in classification could not be explained by criminal history, gender, or age.[212] In follow-up research, various academics and COMPAS developer Northpointe assert that the tools are racially neutral because black and white defendants classified as high risk were

---

[207] Eaglin, *supra* note 31, at 214–18; Eaglin, *May the Odds Be (Never) in Minorities' Favor? Breaking Down the Risk-Based Sentencing Divide*, HUFFINGTON POST (Aug. 22, 2014, 12:30 PM), http://www. huffingtonpost.com/jessica-eaglin/may-the-odds-be-never-in-_b_5697651.html. Recently, two scholars disputed the categorization of certain predictive factors as proxies for race. *See* Jennifer L. Skeem & Christopher T. Lowenkamp, *Risk, Race, and Recidivism: Predictive Bias and Disparate Impact*, 54 CRIMINOLOGY 680 (2016). These scholars asserted that, because certain factors like race, education, and employment cannot alone predict recidivism in black people, these factors cannot be proxies. *Id.* at 704. Yet this study misses my point—education and employment disadvantages predict recidivism as defined by the tools, see *id.*, and those factors disproportionately affect minorities. As the study suggests, lack of education or presence of criminal history would equally result in white defendants and black defendants being classified as higher risk. *See* Skeem & Lowenkamp, *supra*, at 704. The issue, as explained in the text above, is that blacks experience these factors disproportionately. This is particularly true when it is combined with prior arrests as a factor to estimate recidivism.

[208] Chesa Boudin, *Children of Incarcerated Parents: The Child's Constitutional Right to the Family Relationship*, 101 J. CRIM. L. & CRIMINOLOGY 77, 81–82 (2011).

[209] Joseph Murray & David P. Farrington, *The Effects of Parental Imprisonment on Children*, 37 CRIME & JUST.: A REV. OF RES. 133, 135 (2008).

[210] *See* Angwin et al., *supra* note 2.

[211] *Id.*

[212] *Id.*

rearrested at equal rates.[213] Even if the false positive rates were different, the true positive rates were the same.[214]

The discrepancy between these assertions relate to conceptions of fairness.[215] It also relates to the construction of the tools. As explained in Part I, risk tool developers often choose to estimate recidivism risk as chance of arrest based upon factors like prior arrest. Using arrest as the measure of recidivism makes it impossible for black defendants not to be classified as high risk with more frequency given that arrest rates differ by race.[216]

If these are the known inputs—criminal history and social factors disadvantaging minorities—and these are the known outcomes of the tools— more frequently classifying black defendants as higher risk compared to white defendants—the remaining question is why? Why do the tools use this information to predict? Beyond the convenience argument addressed below,[217] there are at least two interrelated explanations relating to race.

One explanation is unconscious bias. Developers unconsciously program tools that disadvantage minorities by making design choices that reflect socially accepted structural inequities in society.[218] In other words, "oops." Professor Anupam Chander recently dismissed this possibility in his essay, *The Racist Algorithm*, where he notes that "because much of societal discrimination is subconscious or unconscious, it is less likely to be encoded into automated algorithms than the human decisionmakers that the algorithms replace."[219] Yet, as this Article demonstrates, the humans developing the tools *are*

---

[213] Anthony W. Flores et al., *False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks."*, 80 FED. PROB., Sept. 2016, at 38, 41; WILLIAM DIETERICH ET AL., NORTHPOINTE, INC., COMPAS RISK SCALES: DEMONSTRATING ACCURACY EQUITY AND PREDICTIVE PARITY (2016), http://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070616.pdf.

[214] *See* Flores et al., *supra* note 213; DIETERICH ET AL., *supra* note 213.

[215] *See* Mayson, *supra* note 195; Jon Kleinberg et al., *Inherent Trade-Offs in the Fair Determination of Risk Scores*, CORNELL U. LIBR. 4 (Nov. 17, 2016), https://arxiv.org/pdf/1609.05807.pdf.

[216] Alexandra Chouldechova, *Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments*, CORNELL U. LIBR. (Feb. 28, 2017), https://arxiv.org/pdf/1703.00056.pdf ; 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.

[217] *See infra* Section II.D.

[218] Selection of predictive variables inevitably disadvantage different groups more. *See* Barocas & Selbst, *supra* note 14, at 688.

[219] Chander, *supra* note 28, at 1028.

decisionmakers. Whether to predict arrests versus convictions, and whether to include certain predictive factors to determine that outcome are choices that developers make.[220]

That developers must explain their decisions internally provides little solace for this threat.[221] Why not? First, developers have their own set of incentives that shape their decisions, and none of those incentives circle around equality. Developers want to construct actuarial tools using cheap, accessible data, even if the data reflects racial biases.[222] This explains the decision to use arrest data as both a predictor and an outcome. Second, developers are tasked with creating risk tools that predict with accuracy. They are not tasked with developing race-neutral tools.[223] To the extent that equality flows from the tool, it is a side effect and not a priority. Third, developers do not permit that the internal explanations throughout the construction process be reviewable by the public. Most developers select risk factors without the opportunity for public review.[224] Yet without the opportunity to review, there is reason to believe their decisions will not further equality principles.

A second explanation follows from the first: developers subordinate the quest for equality in pursuit of accuracy. If society could only have one— equality or accuracy—which would it pick? Empirical debates currently struggle with this question. A tool may accurately estimate risk and consistently classify more minorities as higher risk even though they will not engage in future criminal behavior. Whether society accepts that black defendants will disproportionately bear the burden of additional supervision flowing from actuarial risk assessments is a normative decision. More consideration of this question is beyond the scope of this Article.[225] The point

---

[220] Except for machine learning tools, where developers do not pre-identify risk factors.

[221] *But see* Chander, *supra* note 28, at 1029 ("Because of a programming process that requires both writing down explicit instructions and documenting what particular code does, unconscious or subconscious discrimination is less likely to manifest itself in computer programming than in human decisionmaking.").

[222] This may amount to "rational racism" as well. Rational racism occurs when developers rely upon more simplified data because more granular data that would explain variations would be costlier or challenging to use. *See* Barocas & Selbst, *supra* note 14, at 690; *see also* FREDERICK SCHAUER, PROFILES, PROBABILITIES, AND STEREOTYPES (2003).

[223] *See, e.g.*, INTERIM REPORT 2, *supra* note 65, at 1 (restating the recidivism risk study's goals and nowhere representing an interest in racial inequities); *see also* NORTHPOINTE, INC., *supra* note 45, at 26 (stating its goal to develop tools that predict recidivism). *See generally supra* Section II.A.

[224] As a rare exception, consider Pennsylvania's development of risk tools discussed *infra* Part III.

[225] For what it is worth, the answer to this question is not as obvious as it might appear. The criminal justice system may not be able to bear a decision to sacrifice equality to accuracy. Erin Murphy, *Relative Doubt: Familial Searches of DNA Databases*, 109 MICH. L. REV. 291, 321–22 (2010) (arguing that the

here is far more simple: to the extent risk assessment tools place equality and accuracy at odds with one another, the public should have some input into whether and how to resolve these dilemmas long before the tools have been adopted in a jurisdiction. Input and accountability are necessary during the tool construction process.

In summary, tool developers make decisions during the construction process that impact a defendant's opportunity to benefit from the use of risk assessment tools at sentencing. If these tools are used in the criminal justice system, developers' choices and their effect on equality should be addressed head on and in public.

C. The Purpose of Punishment

Judges must determine sentences that further any or all of the normative purposes of punishment—including retribution or the utilitarian goals of deterrence, rehabilitation, and incapacitation.[226] These boil down to two justifications: either the defendant needs punishment because he deserves it or because it will benefit society at large.[227] States ordinarily do not indicate a particular theory that should guide sentencing.[228] They simply state all the goals and urge judges to sentence in pursuit of a purpose. Yet the normative purposes of punishment set the criteria to determine if a sentence is appropriate.[229] These purposes also set the criteria to determine if a risk tool provides "good" information for sentencing.

Of the myriad rationales for punishment that might underpin sentencing in general, the incapacitation rationale most squarely justifies the use of predictive evidence.[230] In other words, sentencing should aim to prevent social harm. In this sense, actuarial risk tools build upon numerous criminal justice reforms adopted in recent decades—such as habitual offender sentencing enhancements and mandatory minimum penalties—aimed to shape the justice

---

appearance of racial bias in familial DNA searches may undermine the legitimacy of the criminal justice system). The legitimacy of a system that perpetually incarcerates and even kills black men disproportionately has been put into question by leading scholars and, more recently, the #BlackLivesMatter movement.

[226] See, e.g., Aya Gruber, A Distributive Theory of Criminal Law, 52 WM. & MARY L. REV. 1, 4 (2010).

[227] See id.

[228] See RICHARD S. FRASE, JUST SENTENCING: PRINCIPLES AND PROCEDURES FOR A WORKABLE SYSTEM (2013).

[229] Michael Tonry, Purposes and Functions of Sentencing, 34 CRIME & JUST. 1, 10 (2006).

[230] Retribution, as compared to the utilitarian goals, seeks to punish an individual based on moral desert and the defendant's previous wrongdoing. It looks to the past, while risk tools look to the future. Robinson, supra note 35.

system toward preventive detention.[231] This rationale supports incapacitation of a defendant for as long as she presents a risk of danger to society.

Yet current risk tools poorly vindicate that rationale. Advocates endorse risk-based sentencing because these recidivism risk tools provide information regarding whether a defendant poses a risk to society. Such information appears valuable in determining a sentence that furthers incapacitation goals. But the question that courts need to know in relation to incapacitation theory is slightly different: what effect will a particular sentence have on the defendant's likelihood of threatening public safety in the future?[232] This question engages a deeper question about prevention, not simply prediction. Not one predictive tool comes close to addressing this question.[233] Current tools do not consider the length of sentence as a variable on risk. Few tools even consider the offense of conviction that led to sentencing. Currently, tools suggest the probability of recidivism within a set time—usually one to three years—but without any reference to how an actual sentence may affect the defendant's recidivism risk.

Acknowledging that these tools aim to advance preventive detention only further illustrates the difficult normative judgments underlying use of the tools at sentencing. As Professor Michael Tonry explains, "[N]ormative purposes provide the theoretical criteria for deciding whether sentences imposed on individual offenders are just or appropriate."[234] It also indicates what information is pertinent to a sentencing determination. As such, it indicates whether and what predictive tools should measure. These are not empirical questions. As Professor Christopher Slobogin explained in advocating for a preventive detention system, "the degree of risk necessary to authorize intervention . . . are moral/legal questions that laypeople and legal decisionmakers, not clinical experts, should decide."[235] As this Article demonstrates, the very construction of the tool implicates the "moral/legal" questions that Slobogin and others recognize in a shift towards preventive detention. Tool developers are not equipped or situated to address these questions. Only society can decide what tools should predict and why at sentencing.

---

[231] *Id.* at 1438.

[232] *See* HARCOURT, *supra* note 1, at 122–36; Starr, *supra* note 1, at 855–58.

[233] *See* Sonja B. Starr, *The New Profiling: Why Punishing Based on Poverty and Identity Is Unconstitutional and Wrong*, 27 FED. SENT'G REP. 229, 233 (2015).

[234] Tonry, *supra* note 229, at 11.

[235] Slobogin, *supra* note 25, at 167.

D. Developers' Incentives

To complicate the normative judgments identified above, it is important to note that tool developers have incentives in the tool-construction process that diverge from the interests of society at large. Developers construct recidivism risk tools with two sets of concerns in mind: available data and sufficiently varied observations. These interests influence the assumptions and policy decisions embedded in the tools such that tool construction may conflict with or contradict existing sentencing law and policies.

Developers shape the basics of the tool at the outset of the design process based upon available data. Whether data is already available, or how much time and money it may cost to obtain data, will shape tool-construction decisions. For example, because developers wanted to develop the ORAS within three years, they chose to use arrests as the indicator of recidivism, and only studied the likelihood of this event occurring within a year of the initial observation.[236] Using arrest data was simple, low cost, and easy to access. [237] Setting up the study within a short period of time also controlled upfront costs. Consideration of these resource constraints limited the amount of data collected (studying over one year as opposed to more), and shaped the predictive questions that the tool would address. This example demonstrates a broader principle applicable across tool creators—use available data where possible to construct the risk tool.[238]

Developers need sufficiently varied observations in the data sets to effectively predict specific outcomes. Prediction is challenging to achieve unless a data set contains enough people who engage in the behavior that developers wish to predict. For example, imagine a data set that contains only forty-five people who actually committed murder in an eighteen-month period out of 10,000 individuals observed. If a developer wishes to predict homicide, it will be difficult to glean much meaning from the predictors identified in the set.[239] Additionally, the false-negative rate could be very high. If the tool

---

[236] See LATESSA ET AL., supra note 63, at 15.

[237] See Jain, supra note 198, at 818 (stating that "arrest rates are relatively high, making arrests a valuable source of data"); Murphy, supra note 44, at 510–11 (stating that criminal records are easily accessible for data use).

[238] See, e.g., INTERIM REPORT 2, supra note 65, at 1–2 (using data collected for sentencing commission and arrest data to develop risk tool).

[239] This example derives loosely from a proposed data set recently set forth by Dr. Richard Berk. See BERK, supra note 41, at 4–5. The difficulty in predicting violent crime, particularly due to low base rates, is well documented. See Markus Breitenbach et al., Creating Risk-Scores in Very Imbalanced Datasets:

predicts no event of recidivism 98% of the time, it would be highly accurate; however, if someone does commit homicide then the cost to society for the false negative would be high as well, because public safety would be compromised.

To address this concern, developers make choices regarding how to flatten out the data set. In this example, developers may choose to predict something else that will increase the base rate, such as all violent offenses or even all arrests, regardless of the type of offense alleged. In other words, they choose how to simplify the prediction question so that more individuals fit the criteria necessary to achieve variety in the data set sufficient for study and analysis.

The decision to flatten out the data by changing the outcome of interest is problematic because it reflects a normative judgment about what matters at sentencing.[240] Developers may manipulate the data by aggregating distinct types of offenses to produce outcome-significant events of interest in the data set. For example, the VRAG includes any criminal behavior, even if the behavior did not result in an arrest or conviction.[241] The ORAS, on the other hand, uses arrests for any type of offense, including technical violations of probation.[242] Aggregated events have different criminal justice significance. Someone who commits a misdemeanor or technical violation of probation does not pose the same threat to society as someone who will commit a violent crime. However, by flattening out the data to include both types of contact with the criminal justice system as events of interest in a tool, data researchers can easily increase the size and variety in their data set. This manipulation facilitates creating more predictive models. But these events have varying significance,[243] and their occurrence may influence a judge's ultimate sentence

Predicting Extremely Violent Crime Among Criminal Offenders Following Release from Prison, in RARE ASSOCIATION RULE MINING AND KNOWLEDGE DISCOVERY: TECHNOLOGIES FOR INFREQUENT AND CRITICAL EVENT DETECTION 242 (Yun Sing Koh & Nathan Rountree eds., 2010) (noting that "events of interest" occur in less than 20% of participants in violent recidivism studies). Although outside the context of adult recidivism risk tools used at sentencing, the pretrial risk assessment tool in the ORAS system faced a similar problem— too few defendants committed events of interest (failure to appear pretrial). There, Dr. Edward Latessa and his team infused the underlying data set with information on defendants from out-of-state, in which failure to appear was more prevalent. *See* LATESSA ET AL., *supra* note 63, at 14.

[240] As Professor Richard Berk explains, "[T]he choice of what to forecast is a blend of legal, political, and technical concerns." BERK, *supra* note 41, at 10.

[241] *See supra* notes 89–93.

[242] *See supra* note 94.

[243] *See supra* notes 83–88.

differently.[244] Few tools parse out these distinctions with much clarity. The ORAS creators recently developed a model within its original model that predicts misdemeanor offenses specifically for this exact reason.[245] Few other tools have created similarly specific tools.

As a second example, developers rely on unadjudicated behavior—meaning arrests—as an event of interest or a predictive factor given their unique and diverging interests in data size and variety. Currently, predictive tools rely heavily on criminal history events both as variables to predict recidivism and as the events of interest that define recidivism.[246] Tool creators frequently include arrests, charges, and instances of contact with the criminal justice system because they are available. Arrests provide a cheap, easy, and accessible data set for researchers to pull information.[247] Dr. Edward Latessa and his team explicitly chose to use arrests as the recidivism event because the data was readily available and the time frame for producing the tool was short.[248] The Pennsylvania tool uses arrests in part because that data is available from the state database.[249] Using arrests and other complaints about potentially criminal behavior makes the predictive models easier to create cheaply and quickly.

Yet the decision to rely on arrest data rather than convictions has significant meaning at sentencing as well. While sentencing law traditionally permits consideration of a wealth of information with very few limitations,

---

[244] Take as an example recidivism based on parole violation. As the Marshall Project recently explained, "[I]n the current era of criminal justice reform, states have differed in their attempts to incarcerate fewer technical violators. Some have done nothing, while others are implementing a variety of less punitive sanctions for parolees or capping the number of days they can be incarcerated for." Eli Hager, *At Least 61,000 Nationwide Are in Prison for Minor Parole Violations*, THE MARSHALL PROJECT (Apr. 23, 2017, 10:00 PM), https://www.themarshallproject.org/2017/04/23/at-least-61-000-nationwide-are-in-prison-for-minor-parole-violations#.UmaCqOQtq. For more discussion of these alternative sanctions, see Eaglin, *supra* note 37. The point here is that recidivism risk for some specific events, like technical violations of parole, will not carry the same significance at sentencing as others, like risk of violent assault.

[245] Dr. Edward Latessa and his team of data scientists created a tool that predicts misdemeanor offenses due to requests by judges for clarity and nuance in predictive tool outcomes. *See* LATESSA ET AL., *supra* note 56.

[246] *See supra* Part I. For example, the LSI-R uses parole and probation revocations and COMPAS uses "other" supervision violations. *See* Hamilton, *supra* note 57, at 98, 104. The Federal Post Conviction Risk Assessment Scoring Guide, currently not used for sentencing but held out as a particularly accurate tool, "[c]ount[s] all contact with law enforcement resulting from criminal conduct or status offenses (truancy, curfew violations, run-away)." *Id.* at 104 n.145. It also "[c]ount[s] arrests and referrals to court for all offenses (including traffic)," as derived from the official records. *Id.*

[247] *See* Jain, *supra* note 198, at 818; Murphy, *supra* note 44, at 510–11.

[248] *See* LATESSA ET AL., *supra* note 63, at 15–16.

[249] *See* INTERIM REPORT 2, *supra* note 65, at 1.

states vary in where they draw the line on the unadjudicated conduct that may
be considered. There is a strong sentencing policy argument against using
unadjudicated conduct.[250] Some states adhere to this perspective. Either by
legislation or by judicial decision, these states exclude some types of
unadjudicated behavior from consideration at sentencing. Two states that
participate in risk-based sentencing—Minnesota and Washington—preclude
the use of non-conviction offenses as sentencing considerations.[251] Indiana and
North Carolina prohibit the use of acquittals at sentencing.[252] Despite these
prohibitions, two of the four states permit risk-based sentencing and the use of
actuarial tools that do consider these factors.[253] When used at sentencing, such
risk tools directly undermine the state's sentencing policy based on tool-
construction decisions.

   This Part illustrates that tool developers' interests in data size and variety
shapes construction choices that relate to sentencing policy. These interests
shape how entities developing tools conclude the normative judgments
identified in the previous sections. The diverging interests further illustrate that
society, not the tool developers, must independently decide the normative
judgments embedded in tool construction.

---

[250] For example, Professor Kevin Reitz argues that using unadjudicated conduct undermines the
procedural and substantive guarantees of the criminal justice system. *See* Reitz, *supra* note 110, at 548–53.

[251] *See id.* at 535.

[252] *See id.* at 533, 533 n.63.

[253] Indiana developed the IRAS tool, but permits use of other risk assessment tools including the LSI-R at
sentencing. LSI-R uses prior arrests as a predictive factor at sentencing. *See* Hamilton, *supra* note 57, at 94
(explaining that the LSI-R uses prior convictions and prior arrests at sentencing). IRAS-CST uses arrests under
the age of eighteen. *See* UNIV. OF CINCINNATI, *supra* note 51, at 2–4. In Minnesota, supervision agencies use
risk assessment tools like the LSI-R in making sentencing/disposition recommendations. *See* MINN. DEP'T OF
CORR., STUDY OF EVIDENCE-BASED PRACTICES IN MINNESOTA: 2011 REPORT TO THE LEGISLATURE 5 (Dec.
2010), https://www.leg.state.mn.us/docs/2013/mandated/130241.pdf. The Washington State Department of
Corrections developed an actuarial risk tool that considers convictions, not arrests. *See* WASH. STATE INST. FOR
PUB. POLICY, WASHINGTON'S OFFENDER ACCOUNTABILITY ACT: DEPARTMENT OF CORRECTIONS' STATIC RISK
INSTRUMENT 2 (Oct. 17, 2008), http://www.wsipp.wa.gov/ReportFile/977/Wsipp_Washingtons-Offender-
Accountability-Act-Department-of-Corrections-Static-Risk-Instrument_Full-Report-Updated-October-
2008.pdf. North Carolina considered a risk tool for sentencing purposes, but chose not to endorse its use at
sentencing. N.C. SENTENCING & POLICY ADVISORY COMM'N, RESEARCH FINDINGS AND POLICY
RECOMMENDATIONS FROM THE CORRECTIONAL PROGRAM EVALUATIONS, 2000–2008 25 (2009), http://www.
nccourts.org/Courts/CRS/Councils/spac/Documents/correctionalevaluation_0209.pdf. The Commission
endorsed the use of risk assessments at other discretionary stages leading up to or after sentencing, including
the development of "sentencing plans." *Id.* at 15.

III. TOWARD ACCOUNTABILITY IN TOOL CONSTRUCTION

This Article demonstrates that the value of a recidivism risk tool for sentencing is critically connected to key information about the tool's development, including the data set, the factors, the model, and the tool's translation into qualitative risk categories. These construction decisions implicate larger normative questions best left for criminal justice experts and the political process to resolve. Yet tool construction is obscure. Developers may refuse to disclose some or all of the information discussed above, and rarely will they engage in the larger questions at play when the tools are applied to the criminal justice system. Even when they do, there is reason to doubt that tool developers—whether a public or private entity—are well positioned to address the normative questions implicated in tool construction.

As a solution, government entities and tool developers should adhere to various accountability measures in the construction of actuarial risk tools. Accountability here means more than simply ensuring that the tools do what they say they do, as most computer science scholars use the term.[254] To address the threats identified in this Article, risk tools used for the administration of criminal justice must reflect the values of the communities where the tools are applied. This requires engagement with the construction of the tools at various stages of production to ensure that the data-driven outcomes reflect legal and social values when used in the criminal justice system.

The following sections expand upon this call for accountability in tool construction. Section A disentangles accountability from transparency and develops the two kinds of accountability literatures applicable to risk tools in the criminal justice system. Data researchers call for accountability to ensure that tools do what they say they do. Critics concerned about the fairness of risk tools should shift toward demands of accountability as well—to ensure that the tools accord with a community's values. Section B offers a framework to understand various levels of opacity in tool construction: transparency, accessibility, and interpretability. It also suggests interventions necessary at each level to promote the construction of democratically accountable recidivism risk predictions.

---

[254] *See, e.g.*, Chander, *supra* note 28; Kroll et al., *supra* note 28.

A. From Obscurity to Accountability

Obscurity in tool construction is a pressing problem. Many tool developers refuse to disclose some or all of the key information critical to understanding the value of the risk estimates produced by predictive risk tools.[255] Unlike humans, the tools provide no explanation for their results other than the numerical outcomes translated into risk scores.[256] This obscurity produces two kinds of anxieties as tools expand into the administration of criminal justice: do the tools do what they say they do, and are those tools fair? Recent responses to the former concern—do the tools do what they say—call for accountability in tool construction. To address the latter concern about fairness built out in Part II of this Article requires shifting toward accountability as well. To ensure fair construction of risk tools, government agencies and tool developers should create democratic accountability measures that invite the public to engage in the tool-construction and selection process.

The traditional response to obscurity is transparency, but transparency only scratches the surface of the threats risk tools pose to the administration of justice. The limits of transparency are best illustrated through example. Imagine if all tool creators released the data that they collected to predict recidivism. They may even explain why they collected the information as well. This requirement would be beneficial for future innovation[257] and, technically speaking, would bring some transparency to the creation of risk tools. However, to the judge considering how to weigh this information at sentencing, the hard data would be useless. To the defendant classified as high risk, the data would similarly mean nothing. Only entities capable of analyzing the data glean benefit from this requirement, and few criminal justice actors or members of the general public are thus situated.[258] Such a requirement could have the opposite result of obscuring tool construction even while making it technically transparent.[259]

---

[255] *See, e.g.*, State v. Loomis, 881 N.W.2d 749, 761 (Wis. 2016) (discussing how COMPAS treats information about specific factors used in the tool and the weight assigned to those factors as trade secrets, and refuses to disclose them); *see also* Wexler, *supra* note 33 (describing claims of trade secrecy by developers of actuarial risk tools).

[256] *See* Kiel Brennan-Marquez, *"Plausible Cause": Explanatory Standards in the Age of Powerful Machines*, 70 VAND. L. REV. 1249 (2017).

[257] *See* Michael Mattioli, *Disclosing Big Data*, 99 MINN. L. REV. 535 (2014).

[258] Resources are a serious impediment to the expansion of actuarial risk tools as it is.

[259] This requirement could limit understanding of the tool without additional steps to give the outcomes more meaning. The data could be valuable to the court and the defendant if the defense attorney can analyze and test the information. *See* Wexler, *supra* note 33 (arguing that all information about machine tools used in

Transparency alone does not provide insight as to whether the tool's outcomes provide valuable information for use at sentencing. For example, does a tool use information prohibited from consideration at sentencing under state law? Does the data set reflect individuals from the communities where the tools are being used? Did tool developers include factors that will disproportionately impact the risk scores of poor and minority defendants? If so, which ones? Perhaps the better question is why? Are the cut-off points located in line with public's judgment about how much risk is the appropriate amount to accept in society? We often do not know. Information about the data set or even the weighting of various factors does not, on its face, provide insight to these questions either. Yet answering these questions is critical to understanding and giving value to the tool's result.

To be sure, transparency has an essential place in ensuring fairness in automated prediction tools, and data scholars often urge its expansion. [260] A growing body of big-data scholarship, however, urges accountability rather than simply transparency.[261] While recognizing that accountability means ensuring administrators "choose the approach that ... works best for their communities,"[262] the focus is largely on ensuring that tools do not controvert those goals through technology. Accountability measures, including the development of additional computer programs to check system function, can achieve this and more.[263] This call for accountability has significance in the criminal justice context as well, as it is critical that courts can rely on tools that do what they say they do.[264] Solutions include auditing the data collection process and issuing training programs for courts and administrators to ensure the proper application of the technique for sentencing.[265]

Yet in the context of criminal justice, accountability takes on a greater meaning as well: do the actions of the government reflect the values of the

---

the criminal justice system should be transparent so that lawyers can educate the courts about their fallibility). I do not suggest here that such interventions are not meaningful. Still, this Article aims to provide measures that give risk assessments meaning for the broader public. This meaning complements and precedes individualized interjections.

[260] See, e.g., PASQUALE, supra note 15; Citron & Pasquale, supra note 15, at 6–8; Kroll et al., supra note 28.

[261] See Chander, supra note 28; Kroll et al., supra note 28.

[262] Ferguson, supra note 12, at 58; see also Tal Z. Zarsky, Transparent Predictions, 2013 U. ILL. L. REV. 1503, 1533 (2013).

[263] See Kroll et al., supra note 28.

[264] See, e.g., Ferguson, supra note 12.

[265] Id. (discussing implementation problems in the policing context).

people? In the context of actuarial risk tools, do the risk estimates reflect whether and how society wants to assess a defendant's risk level at sentencing? Such accountability varies from the kind called for by data scholars. A predictive tool could estimate a defendant's level of risk just as it says it does without technically considering race or other impermissible factors, but nevertheless the result is illegitimate because it frustrates societal values of accuracy or equality in the criminal justice system. While one could ask an engineer or a developer about how the tool works, that developer's value judgment may not be consistent with the community where the tool is being applied. The community of application is the relevant metric by which to measure recidivism risk. Validity studies provide no insight as to that question. At times, bureaucratic officials will not be able to do so either.[266] Only the communities affected by the tools can voice those values. When automated predictive tools are applied in the administration of criminal justice, nothing less than democratically accountable recidivism risk predictions should suffice.

Democratic participation in the construction of actuarial risk tools is essential when the tools are developed to facilitate administering criminal justice. Risk tools represent a form of evidence envisioned to inform criminal justice actors about the defendant's risk level in relation to society at large. At times, the information is produced with the sanction of the state, like when the state works with the developers to construct the tool.[267] More often, however, the tools are developed by private entities and adopted by jurisdictions with limited opportunity for expert input and localized feedback.[268] Yet constructing recidivism risk is not an objective endeavor; rather, it is laced with "profound policy questions that must be resolved in democratically accountable ways."[269] Who gets to make those decisions and when are critical to determining whether that information accurately reflects the values of the community of application. Lawmakers and policymakers should create interventions that pierce the perceived objectivity of risk tools and facilitate engagement with the underlying normative judgments implicated through construction. Such interventions, including some suggestions discussed below, are critical to ensuring that normative choices embedded in the tools reflect society's judgments about what counts for recidivism risk at sentencing, how, and why.

---

[266] *See* Stephanos Bibas, *Transparency and Participation in Criminal Procedure*, 81 N.Y.U. L. REV. 911 (2006) (describing the tension between bureaucratic "insiders" like judges, police, and prosecutors versus "outsiders" like crime victims, bystanders, and the general public).

[267] *See supra* notes 47–49.

[268] *See supra* notes 44–46.

[269] Barry Friedman & Maria Ponomarenko, *Democratic Policing*, 90 N.Y.U. L. REV. 1827, 1836 (2015).

There are costs to infusing the tool-construction process with criminal justice expertise and political process accountability. To start, it will slow the process of developing risk tools. This means fewer tools will become available in the foreseeable future. That is not necessarily a bad thing, as the tools are developing without much caution. The criminal justice system has survived a long time without predictive tools. Slowing down creation and adoption of tools can promote reflection about their construction and use. It can also prompt society to grapple with the underlying normative challenges that the tools present.

Infusing this kind of accountability into the construction of recidivism risk tools will present implementation challenges. How can developers solicit the requisite input in accountable ways? There are a variety of approaches, some of which states and local jurisdictions have already adopted. As one example, public notice and comment on normative decisions throughout the development process would resolve many of the issues raised here.[270] Requiring jurisdictions to hold public hearings and vote on the selection of various risk tools or specific features of a risk tool is another option. Creating legislation that intervenes with the construction of risk tools, for example defining recidivism risk for predictive tools or preventing the transfer of public data to private ownership, is another possibility. Future research can address the implementation problem in more detail. The point here is to emphasize that measures are necessary to infuse accountability into the construction process.

Likely the biggest opposition to this approach lies in concern over accuracy. Opponents may suggest that infusing accountability into the construction of risk tools could undermine the predictive accuracy of the tools' results. In one sense it may, as the tools may not be constructed as developers wish them to be. On the other hand, if the tools produce more fair and accurate information about what risk means in society, then the tools better fulfill their promise to inform criminal justice decision makers about the defendants in relation to their communities. If pursuing this kind of accuracy is a cost, it is well worth incurring to ensure the quality of the information used to administer justice.

---

[270] For more on the implementation of a notice-and-comment process as applied to criminal justice policymaking, see Richard A. Bierschbach & Stephanos Bibas, *Notice-and-Comment Sentencing*, 97 MINN. L. REV. 1 (2012).

The benefits to this approach are high as well. The criminal justice system is already opaque to most laypeople entering or observing it.[271] Risk tools threaten to make the criminal justice system even more opaque, as the defendant, the judge, and the general public may not understand why someone is considered to have a low, medium, or high risk of recidivism.[272] Shifting normative judgments about tool construction toward the public arena could make society more aware of how the system works. It may increase the system's legitimacy in the public's eye. At the very least, it will produce more clarity about what the recidivism risk estimates mean.

B. Constructing Democratically Accountable Risk

Infusing criminal justice expertise and political process accountability into the construction of actuarial risk tools requires a framework to understand the existing levels of opacity that people need to break through to facilitate meaningful participation. This section provides that framework and suggests interventions at each level.

There are three levels of opacity in the construction of recidivism risk tools.[273] The first level is transparency. Although this Article urges a shift toward accountability, transparency about the tool's design and its use are necessary components. The second level of opacity refers to the issue of accessibility. The communities considering a tool must understand the tools enough to engage with the normative construction choices through the political process. Third, and finally, is the matter of interpretability. Why a tool produces the results that it does and the way its results will be used is critical to

---

[271] *See* STEPHANOS BIBAS, THE MACHINERY OF CRIMINAL JUSTICE 34–38 (2012) (describing the opacity of the criminal justice system); Jocelyn Simonson, *The Criminal Court Audience in a Post-Trial World*, 127 HARV. L. REV. 2173 (2014) (describing the opacity of the criminal justice system to those who attend hearings).

[272] *See, e.g.*, State v. Loomis, 881 N.W.2d 749, 774 (Wis. 2016) (Abrahamson, J., concurring) (noting the Wisconsin Supreme Court's "lack of understanding" as a "significant problem" to understanding a risk assessment tool); Brief for the Public Defender of Indiana as Amicus Curiae Supporting Petitioner at 8, Malenchik v. Indiana, 928 N.E.2d 564 (Ind. 2010) (No. 79S02-0908-CR-365) (noting that counsel for a convicted person will have to "ferret[] out" information about what high risk means for a tool and what it means in the context of setting sentences).

[273] This framework draws upon the insightful framework proposed by Professor Jenna Burrell to clarify the layers of opacity in machine learning algorithms. *See* Jenna Burrell, *How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms*, BIG DATA & SOC'Y, Jan.–June 2016, at 1. Here, I use the terminology of Burrell's framework, but in service to a unique and largely ignored aim: to engage the public in the normative debates about the construction of risk assessment tools used at sentencing. This framework has the benefit of applying to current non-machine learning tools at sentencing and potentially applying to future machine learning tools as well.

understanding construction choices. Each level triggers different questions about the tools and the justice system. The purpose here is to identify the issues in hopes that future research will pursue solutions as predictive tools continue to expand in the administration of criminal justice. Still, at various points this Article provides examples from different states and jurisdictions that have addressed these questions head on.

1. Transparency Measures

Transparency is a necessary step to accountability.[274] Without placing the matter of tool construction in the public spotlight, it would be impossible for a community to engage with the underlying normative judgments implicated throughout the process. Transparency in this context has two separate meanings to ensure democratically accountable risk estimates and temper the threats to sentencing law and policy identified in this Article.

Tool Developers. For tool developers, some transparency is necessary to ensure that localized criminal justice experts can provide input regarding the construction and adoption of risk tools in a jurisdiction. For example, information about the specific origin of the data set underlying a tool is an important disclosure to determine the value of a tool. In addition, the selection of risk factors should be disclosed to ensure that a tool does not contradict the state's existing sentencing law and policy. This information need not include public release of the tool's algorithm. However, even these minimal disclosures go beyond what many developers currently provide.

Tool creators, particularly private entities, face strong disincentives from sharing information about normative policy choices made during the tool-construction process. There is little incentive to disclose data set choices given the emphasis on validity studies as the indicator of tool quality.[275] There is no economic incentive to disclose these choices either.[276] Competition amongst tool creators to develop commercially viable risk assessment tools encourages developers to remain vague about the subjective judgments embedded in their tools. Disclosing specific information about tool-construction choices may lead a consumer to perceive the underlying data set as methodologically weak or

---

[274] Zarsky, *supra* note 262, at 1533–34 ("Transparency is an essential tool for facilitating accountability because it subjects politicians and bureaucrats to the public spotlight.").

[275] *See supra* note 193. Trade secrecy creates another disincentive. For more discussion, see *infra* notes 292–293.

[276] *See* Mattioli, *supra* note 257.

unsound, and ultimately seek out another product.[277] One can imagine a competitor exploiting any weaknesses disclosed and using that to persuade a state or specific jurisdiction toward adoption of its alternative tool.[278] Even when products are soundly developed, detailed disclosure of tool design permits replication of technique, thus allowing new competitors into a developing market.[279]

These disincentives explain the systemic difficulty in obtaining information about the design, development, and evaluation of privately created risk assessment tools used in the criminal justice system. In 2015, University of Maryland School of Journalism Professor Nicholas Diakopoulos undertook a semester-long project to obtain "documents, mathematical descriptions, data, validation assessments, contracts, and source code" related to actuarial risk assessment tools used in the fifty states for any criminal justice determination, including parole, probation, bail, or sentencing.[280] After submitting formal requests to the government agencies using these tools, few states provided actual insight. As Professor Diakopoulos explains, nine states refused the request because a private company owned the information.[281] In essence, the companies treat such design information as trade secrets protected by intellectual property laws.[282] States that entered into private contracts with nonprofit organizations similarly refused to release information about tool design because contractual provisions prevented disclosure.[283]

---

[277] *See id.* at 549; Murphy, *supra* note 44, at 536 (explaining private sector industries' incentive to market their tools).

[278] *See* Mattioli, *supra* note 257, at 549.

[279] As Professor Michael Mattioli notes, "[M]ost big data products cannot be reverse-engineered to reveal the processes that went into their creation" because it is near impossible to "guess the various techniques and judgments that go into processing a dataset." *Id.* at 573, 573 n.171. Given this reality, tool creators' disclosure of tool design is the only way to understand the subjective policy choices embedded in the tool. This is the only way for an outsider to challenge the reliability of the underlying data set, too.

[280] *See* Nicholas Diakopoulos, *We Need to Know the Algorithms the Government Uses to Make Important Decisions About Us*, CONVERSATION (May 23, 2016, 8:48 PM), http://theconversation.com/we-need-to-know-the-algorithms-the-government-uses-to-make-important-decisions-about-us-57869?utm_medium=email&utm_campaign=Latest%20from%20The%20Conversation%20for%20May%2023%202016%20-%204912&utm_content=Latest%20from%20The%20Conversation%20for%20May%2023%202016%20-%204912+CID_efe310bf05b2dc19249223110c254baf&utm_source=campaign_monitor_us&utm_term=he%20writes.

[281] *Id.*

[282] *See, e.g.*, State v. Loomis, 881 N.W.2d 749, 761 (Wis. 2016) ("Northpointe, Inc. . . . considers COMPAS a proprietary instrument and a trade secret."). For more discussion on the intersection of trade secrecy laws and big data, see, for example, PASQUALE, *supra* note 15, at 12–14; Mattioli, *supra* note 257, at 550–56. For a discussion of its application in the criminal justice context, see generally Wexler, *supra* note 33.

[283] For example, Kentucky refused to disclose information in response to the journalist's request for this reason. *See* Diakopoulos, *supra* note 280.

Public entities developing risk tools also face disincentives that support opacity in tool creation. Disclosing elements of a predictive model may encourage strategic behavior by individuals who recognize themselves as low risk.[284] It may be onerous and overly technical, thus providing little insight into the policy decisions at all.[285] Additionally, the result of such disclosures may be a less precise tool. The Pennsylvania Commission on Sentencing's experience in tool development illustrates this dilemma. During its study, the Commission found that the location from which defendants originate was a strong predictor of recidivism.[286] Data researchers initially decided to include this measure in the tool.[287] It was only after public backlash against this predictive factor that tool creators pulled that information from the predictive model.[288] Tool creators perceived this decision as undermining the tools' ability to predict accurately.[289] As such, the public intervention conflicted with the interests of the state agency.[290]

In response to these disincentives, state or local government bodies could create statutes or ordinances that require specific disclosures if the tools are used for the administration of criminal justice. These requirements may run up against companies that seek to protect risk-tool-construction information as commercially valuable, and thus a trade secret that cannot be disclosed. Whether such claims have merit in the criminal justice context is a matter of debate.[291] Still, preemptory transparency requirements can avoid dilemmas that trade secret claims may present. Developers can disclose most tool-construction choices without disclosing the actual algorithm to the public.[292]

---

[284] *See* Kroll et al., *supra* note 28, at 658.

[285] *See id.* at 659–60

[286] INTERIM REPORT 3, *supra* note 139, at 6.

[287] *See id.*

[288] PA. COMM'N ON SENTENCING, PROPOSALS PUBLISHED IN PENNSYLVANIA BULLETIN: ANNEX B (2017), http://pcs.la.psu.edu/guidelines/proposed-for-public-comment-sentence-risk-assessment-instrument/annex-b/view.

[289] PA. COMM'N ON SENTENCING, RISK/NEEDS ASSESSMENT PROJECT: SPECIAL REPORT: IMPACT OF REMOVING DEMOGRAPHIC FACTORS 1 (2015), http://pcs.la.psu.edu/publications-and-research/research-and-evaluation-reports/risk-assessment/phase-ii-reports/special-report-impact-of-removing-demographic-factors/view.

[290] The agency actually recommended that the Commission keep all demographic factors, including county of origin. *Id.* Public pressure explains the decision to ultimately remove the factor. *See infra* Section III.B.

[291] *See, e.g.*, David S. Levine, *Secrecy and Unaccountability: Trade Secrets in Our Public Infrastructure*, 59 FLA. L. REV. 135, 140 (2007) (questioning the applicability of trade secrecy when private companies operate in public infrastructures); Wexler, *supra* note 33 (demonstrating the uncertain application of trade secrecy in the criminal context).

[292] *See* Selbst, *supra* note 66.

Moreover, a jurisdiction can demand the information upfront, before deciding whether to adopt a tool.[293] The decision not to disclose specific information should be available to the public. That way, a jurisdiction can choose not to adopt a tool that refuses to meet its disclosure requirements.

Criminal Justice Actors. For criminal justice actors, transparency requires visibility and specificity about the adoption of risk-based sentencing practices. Legislators mandate use of these tools in numerous jurisdictions. To name a few, the Kentucky, Ohio, Pennsylvania, and Virginia legislatures have concluded that, at some point in the sentencing process, recidivism risk tools must be considered.[294] Even when use of a risk tool is not required, legislatures in several states have encouraged their use directly through statute. For example, the Washington State legislature encourages consideration of recidivism risk tools at sentencing if available.[295] Similarly, Louisiana permits the use of a validated risk-and-needs tool at sentencing for eligible defendants.[296] Some legislatures indirectly encourage use of actuarial risk tools through financial incentives as well. For example, the Illinois Crime Reduction Act of 2009 provided funding for the award of grants to counties that created standard plans to reduce prison commitments by 25%.[297] Several of these programs require use of LSI-R or another risk assessment tool to select

---

[293] *See* Kroll et al., *supra* note 28, at 665–69 (suggesting methods for developers to make sensitive information available upfront via technology for later disclosure).

[294] *See* ARIZ. CODE OF JUD. ADMIN. § 6-201.01(J)(3) (Westlaw through 2017) ("For all probation eligible cases, presentence reports shall . . . contain case information related to criminogenic risk and needs as documented by the standardized risk assessment and other file and collateral information."); IDAHO CODE ANN. § 19-2517(1) (West Supp. 2015) ("If the court orders a presentence investigation to be conducted, the investigation report shall include current recidivism rates for . . . [specified offenders]."); KY. REV. STAT. ANN. § 532.007(3) (West Supp. 2016) ("Sentencing judges shall consider . . . the results of a defendant's risks and needs assessment included in the presentence investigation . . . ."); OHIO REV. CODE ANN. § 5120.114(A)(1)–(3) (West Supp. 2017) ("The department of rehabilitation and correction shall select a single validated risk assessment tool for adult offenders. This assessment tool shall be used . . . [for sentencing or another purpose] . . . ."); OKLA. STAT. ANN. tit. 22 § 988.18(B) (West Supp. 2011) (requiring any felony offenders considered for community punishment to receive assessment under the LSI or "another assessment and evaluation instrument designed to predict risk to recidivate approved by the Department of Corrections"); 42 PA. STAT. AND CONS. STAT. ANN. § 2154.7(a) (West Supp. 2017) ("The commission shall adopt a sentence risk assessment instrument for the sentencing court to use to help determine the appropriate sentence . . . .").

[295] *See* WASH. REV. CODE ANN. § 9.94A.500(1) (West Supp. 2015) (declaring that the court "may order the department to complete a risk assessment report," and "[i]f available before sentencing, the report shall be provided to the court").

[296] *See* LA. STAT. ANN. § 15:326(A) (2015) (stating that criminal courts "may use a single presentence investigation validated risk and needs assessment tool").

[297] *See* Illinois Crime Reduction Act of 2009, 730 ILL. COMP. STAT. ANN. 190/20 (West Supp. 2016).

defendants eligible for diversion.[298] Policy advocates like the Justice Reinvestment Initiative and the National Institute of Corrections are attributed with spreading the use of risk assessment tools at sentencing across the states as well.[299]

Yet most states engaging in risk-based sentencing do not specify which tools can be used. The American Law Institute calls upon sentencing commissions to develop risk tools for sentencing.[300] As noted above, most states do not develop their own risk tools.[301] More often, legislatures are silent on how to select a risk tool, leaving that determination to other criminal justice actors in specific jurisdictions. For example, correctional departments often de facto determine which risk tools to use at sentencing in some states because the courts just use whichever tools are already in use in the jurisdiction.[302]

Criminal justice actors should transparently disclose the specific tools adopted in a jurisdiction and seek community input on that decision if it chooses to engage in risk-based sentencing. A jurisdiction could specify a tool for use through legislation. Along these lines, the Virginia legislature serves as an example because it directed the state sentencing commission to develop a risk assessment tool.[303] A jurisdiction could pursue the notice-and-comment process before adopting a particular tool. In Pennsylvania, the state's sentencing commission sought public comment about the construction of the actuarial risk tool after the state legislature mandated its creation.[304] It provided public information throughout the tool-construction process, from the development of the statistical model to the construction of the tool.[305] These steps invited democratic participation in some of the normative judgments embedded in the resulting tool. For example, a group of citizens created a risk

---

[298] *See Adult Redeploy Illinois*, WILL COUNTY PUB. DEFENDER (2013), http://www.willcountypublicdefender.com/resources/the-court-process/adult-redeploy-illinois-ari.

[299] *See* CASEY ET AL., *supra* note 7, at 37–38; Eaglin, *supra* note 37, at 609–10 (noting the Justice Reinvestment Initiative's endorsement of using risk and needs assessments at sentencing); Klingele, *supra* note 1, at 566 (attributing to the Justice Reinvestment Initiative, the National Institute of Corrections, and state and local initiatives a critical role in expansion of risk assessment tools at sentencing).

[300] MODEL PENAL CODE: SENTENCING § 6B.09 (AM. LAW INST., Tentative Draft No. 2 2011).

[301] *See supra* notes 45–52; *see also, e.g.*, 42 PA. STAT. AND CONS. STAT. § 2154.7 (West Supp. 2017) (sentencing commission making choice).

[302] *See, e.g.*, OHIO REV. CODE ANN. § 5120.114 (West Supp. 2017) (corrections department making choice); OKLA. STAT. tit. 22 § 988.18(B) (West Supp. 2011) (corrections department making choice).

[303] VA. CODE ANN. § 17.1-803 (West 2013).

[304] *See* PA. COMM'N ON SENTENCING, *supra* note 288.

[305] *See id.*

assessment task force to comment on and track the development of the state's risk tool.[306]

Future scholarship should consider how to infuse more transparency into the construction and adoption of actuarial risk tools across the country. The point here is simply that criminal justice actors and tool developers have an obligation to provide enough transparency about tool construction and selection to facilitate accountability regarding the tool's adoption.

2. Accessibility Measures

The public cannot engage with tool construction without information to help give normative choices meaning. Recidivism risk prediction is, on its face, largely inaccessible to the public. Accessibility demands that those who construct the tools and those who use the tools understand the impact of the estimates in application so as to better inform determinations of whether and how to construct the tools. This demand has intersecting meanings for tool developers and criminal justice actors as well, depending on the measure of intervention.

Defining Recidivism and Inputs. Jurisdictions interested in pursuing the use of actuarial risk tools at sentencing should invite criminal justice expertise and community input regarding the definition of recidivism and whether various factors should be eliminated from tool consideration as a matter of policy. A state agency could publicize options to define recidivism for sentencing or indicate alternative tools that define recidivism differently for public input regarding tool selection. Information about the implications of each recidivism definition should be produced for public consideration.[307] Similar information should be produced regarding the risk factors a tool considers. Criminal justice experts and the jurisdictions where tools may be applied can then use that information to access the debate about whether, when, and how to engage in recidivism risk prediction at sentencing.

---

[306] *See* Marni Jo Snyder, Attorney, Testimony on Behalf of the Risk Assessment Task Force (May 23, 2017), http://www.pahouse.com/files/Documents/2017-05-25_100857__Testimony%20before%20Sentencing%20Commission.pdf.

[307] This intervention aligns with calls to critically engage with the construction of actuarial risk tools in other criminal justice contexts like pretrial bail detention. *See* Gouldin, *supra* note 166 (proposing a study on alternative definitions of flight risk that are more precise to the judicial concerns of pretrial detention determinations).

These interventions can have meaningful results in preventing the construction of tools that are inconsistent with a community's values. For example, the Pennsylvania Commission on Sentencing attempted to use neighborhoods as a predictive factor in its risk assessment tool.[308] Public critique of that factor led the Commission to conduct a study demonstrating the impact of using that factor in a risk tool.[309] It also led to exclusion of county of origin as a predictive factor in the proposed tool.[310] Other factors that disproportionately affect members of particular communities may also be excluded if developers systematically disclose the specific factors used for prediction.

Recidivism Risk Classifications. Regarding the classification of various risk estimates into specific risk categories, input from criminal justice experts and political process engagement are already feasible. Tool developers currently provide opportunities for criminal justice actors to decide the placement of cut-off points and to allocate the costs of classification errors. Criminal justice actors inconsistently take that action.

Government agencies should always decide the cut-off points. In some instances, agencies offer guidance to developers on how to categorize defendants' risk levels. For example, Virginia's legislature provided insight on how many low-level offenders it wanted the risk tool to recommend for diversion (the lowest 25%) through legislation.[311] The Virginia Sentencing Commission then used deciles to group cases of a normative sample into risk bins.[312] In other words, it chose cut-off points based on the number outcome, and subjectively chose how much of the sample population should fit into each category. Entities developing tools are quite receptive to guidance on this issue.[313]

---

[308] PA. COMM'N ON SENTENCING, *supra* note 289, at 1.

[309] *Id.* (citing Starr, *supra* note 32) (noting Starr's article as a motivation to study the impact of demographic factors on the proposed risk tool).

[310] PA. COMM'N ON SENTENCING, *supra* note 288 (predictive factors include age, gender, prior arrest, prior arrest offense type, current conviction offense type, multiple current convictions, prior record score, and prior juvenile adjudication).

[311] Kern & Farrar-Owens, *supra* note 47.

[312] *See* VA. CRIM. SENTENCING COMM'N, ASSESSING RISK AMONG SEX OFFENDERS IN VIRGINIA 92 (Jan. 2001), http://www.vcsc.virginia.gov/sex_off_report.pdf (explaining that the cut-off point is twenty-eight points).

[313] *See, e.g.,* Berk, *supra* note 189, at 1079 (explaining that stakeholders are receptive to selecting cost ratios in context of risk tools used at parole).

Tool developers should not provide default cut-off points. Some developers currently offer automatic default cut-off points for government agencies that do not want to make these decisions. For example, developers of the ORAS tools introduced cut-off points based on their own analysis of the data.[314] Northpointe offers the tool with cut-off points introduced already, although it gives state actors a choice to tinker with this aspect of tool design. If Northpointe introduces the cut-offs, it uses a separate algorithm to divide the sample into decile groups then accorded a risk level.[315] This option should be eliminated. How developers would choose to categorize defendants will be influenced by their own sets of interests. These interests are not necessarily representative of the values of a jurisdiction using a tool.

Governmental entities must engage the communities affected by the tools in this decision to ensure tools reflect normative policy judgments in accessible ways. For example, government agencies could issue statements explaining where they would place cut-off points and why before deciding to adopt a tool for risk-based sentencing. Such statements would facilitate informed political process engagement with that decision.

Outputs. For criminal justice actors and the public to engage with the tools, developers should produce information that facilitates accessibility as well. Tool developers should be forthcoming about the outputs a tool produces to facilitate public accountability measures.[316] For example, what percentage of defendants classified as high risk actually commit a crime in the future? What is the racial and ethnic breakdown of the risk classifications? What kind of socioeconomic impact does a tool-construction decision have?[317] This information is not easily accessible, but it is necessary to ensure tool-construction decisions reflect a community's values.

While tool developers have little incentive to disclose information voluntarily,[318] lawmakers could require that developers produce and publish impact studies as a condition for tool adoption in their jurisdiction. State

---

[314] LATESSA ET AL., *supra* note 63, at 17.

[315] Where possible, each group would have approximately equal-sized numbers of offenders. *See* NORTHPOINTE, INC., *supra* note 45, at 8.

[316] *See* Chander, *supra* note 28, at 1039 ("Instead of *transparency in the design* of the algorithm, what we need is a *transparency of inputs and outputs*.").

[317] *See generally* Starr, *supra* note 1 (noting the constitutional implications of risk tools due to socioeconomic impact). Information for a particular jurisdiction could be valuable before the public provides input on whether to use the tools at sentencing or in some other criminal justice context.

[318] *See supra* Section II.D.

sentencing commissions could produce the information as well through impact statements before tool adoption. Such aggregate information would likely engage public discourse.[319] One need only look to the debate spurred by the ProPublica study on race and risk assessments to demonstrate the point.[320] With the necessary information and opportunity, the public can meaningfully engage in the debate about whether and how tools should produce information for sentencing.

In summary, tool developers and criminal justice actors have an obligation to make recidivism risk construction choices accessible. Widespread and targeted educational efforts are necessary to make the public more knowledgeable about how recidivism risk construction decisions will affect their communities.[321] The meaning of various decisions must be clear so that expert and lay members of the community can provide valuable feedback on how to develop a tool used for the administration of criminal justice in its jurisdiction.

3. Interpretability Measures

Those using the tools must be able to interpret the results as well. Interpretability refers to the "why" questions. Why did a tool produce the results that it did? More importantly, why does actuarial risk assessment fit into the administration of criminal justice, and where? Again, these questions have different meanings for tool developers and criminal justice actors.

Statistical Modeling. As noted above, there are a variety of statistical models that can be adopted to predict recidivism risk. Most tools currently use traditional regression models, but machine learning methods are on the horizon.[322] Risk tools using this modeling create difficult interpretability issues, as the developers creating the tools cannot explain what factors a tool uses to predict recidivism risk.[323] The quickly evolving methods of prediction only further illustrate the need for transparency and accessibility measures that

---

[319] As an example, a Pennsylvania Risk Assessment Task Force now calls upon the Sentencing Commission to publish results concerning the racial impact of tools before adoption of the proposed risk assessment. *See* Snyder, *supra* note 306.

[320] *See supra* Section II.B, notes 210–214.

[321] *See* Burrell, *supra* note 273, at 4.

[322] *See supra* notes 41–43.

[323] *See* Chander, *supra* note 28, at 1040 ("[I]n the era of self-enhancing algorithms, the algorithm's human designers may not fully understand . . . what some of their algorithms do."); *supra* notes 41–42 and accompanying text (describing machine based learning methods).

invite public engagement with tool construction. Even as the methods of prediction change, the underlying normative judgments regarding accuracy, equality, and purpose of punishment will persist. Whether these types of models are appropriate for tools used at sentencing is an open question that future research promises to address.

The Purpose of Punishment. The long-debated question of why we punish is beyond the scope of this Article. Suffice it to say that scholars make sound arguments in favor of each purpose. Unfortunately, most states choose not to select a particular purpose, instead leaving it to individual judges to select a guiding theory to justify a particular sentence. The wisdom of this decision is debated.

Yet insight on the purpose a tool seeks to further in application is necessary for valuable input regarding tool construction. Even if a jurisdiction does not select one guiding purpose of punishment, it should invite input and announce which of the primary purposes a risk tool should further. This would occur before soliciting public input regarding selecting or adopting a tool, which would better inform their input on later construction decisions.

This determination could inform and shape decisions about tool construction and adoption. For example, if a jurisdiction seeks a tool meant for rehabilitation, developers should offer only tools designed to identify particular risks that require interventions available at sentencing in that jurisdiction. If deterrence is the guiding purpose, developers should offer tools that address how much supervision or incarceration would reduce the likelihood of future criminal behavior in that jurisdiction. If these tools are not available or applicable, then the jurisdiction should not pursue risk-based sentencing further.

If a jurisdiction chooses to pursue a risk tool that furthers incapacitation— as many would[324]—this would inform various tool-construction decisions as well. Tools could not be described as a solution to reduce mass incarceration, which may temper some enthusiasm for the reform. Rather, the tools would be described as a mechanism to ferret out anyone who poses a risk to society for the purpose of additional detention or supervision. That aim would further clarify the normative judgments for public input throughout the tool-

---

[324] *See* Eaglin, *supra* note 31, at 222–24 (noting that criminal justice reforms are often motivated by a desire for total incapacitation).

construction or tool-adoption process. For example, it would influence where a jurisdiction locates the cut-off points between risk categories.[325]

Other benefits could flow from this interpretability measure as well. Announcing the tool's purpose could prevent a one-size-fits-all approach to risk assessment tools at sentencing. Because sentencing presents unique limitations regarding the factors that can be considered to determine punishment, risk tools in this context may be very different from those used at other points in the justice system. In line with these determinations, a jurisdiction should prohibit the application of risk tools designed for one purpose to be used for another. Additionally, this announcement could motivate a broader discourse about the purposes sentencing should pursue.

CONCLUSION

Risk-based sentencing seeks to infuse data-driven technology into the determination of punishment through the introduction of actuarial estimates of a defendant's recidivism risk. Although conceived as objective and helpful information, constructing an actuarial risk tool raises longstanding questions about accuracy, equality, and the purpose of punishment that need to be addressed. This Article examines the tool-construction process to bring forth the normative judgments embedded in the tools' development. These judgments concern important values at sentencing too complex and contested to leave in the hands of tool developers alone. This Article calls for democratic accountability measures to address threats that risk-tool construction presents at sentencing. It proposes measures to ensure that a tool's results reflect values consistent with those of the community adopting that tool.

In many ways, this Article raises more questions than answers. How do we balance innovation and accountability? What do courts do with the information produced by risk tools now without necessary measures of accountability?[326] Given the opacity of the judgments entrenched in these tools, should we move in the direction of big-data criminal justice at all? The answers to these questions are unclear, and deserve more discussion in scholarship and the public discourse. This Article recognizes that the answers to these

---

[325] *See also* Mayson, *supra* note 1.
[326] *See* Jessica M. Eaglin, *Technological Evidence and Judicial Sentencing Discretion* (forthcoming 2018).

determinations will not only affect whether courts use the predictive risk tools, but how those tools are constructed.

One point, however, is clear. More caution and nuance is necessary in approaching the use of recidivism risk tools in the administration of criminal justice. Indeed, former Attorney General Eric Holder expressed concern and urged caution in the use of risk assessment tools at sentencing more than two years ago.[327] Although measured in his call, he was criticized harshly.[328] This Article echoes some of the concerns he raised, including the effect risk tools may have on equality at sentencing. It calls for tool developers and criminal justice actors to facilitate more public engagement with that ongoing debate.

---

[327] Holder, *supra* note 32.

[328] *See, e.g.*, Judge Richard George Kopf, *Like the Ostrich that Buries Its Head in the Sand, Mr. Holder Is Wrong about Data-Driven Sentencing*, HERCULES AND THE UMPIRE (Aug. 10, 2014), https://herculesandtheumpire.com/2014/08/10/like-the-ostrich-that-buries-its-head-in-the-sand-mr-holder-is-wrong-about-data-driven-sentencing (criticizing former Attorney General Eric Holder's critique of risk-based sentencing); Sheldon Whitehouse, Letter to the Editor, *Useful Tools in Sentencing*, N.Y. TIMES (Aug. 18, 2014), http://www.nytimes.com/2014/08/19/opinion/useful-tools-in-sentencing.html (arguing that risk assessment tools play an important role in the administration of criminal justice).