

RACE AND RISK ASSESSMENT:

WOULD WE KNOW A FAIR TOOL IF WE SAW IT?

BY SARAH PICARD-FRITSCHÉ, JULIAN ADLER,
YOLAINE MENYARD, AND SHUBHA BALA

With the United States housing 5% of the world's population, yet 25% of its prison population, mass incarceration has emerged as a peculiarly American political problem. Despite growing momentum in favor of reform, incarceration remains the default response to both serious and minor crime in many jurisdictions. Worse, many jurisdictions continue to send people of color to jail at disproportionately higher rates than their white counterparts. While racial disparities in incarceration stem partly from historic inequalities, they are arguably sustained, in some measure, through implicit bias at key decision points such as arrest and sentencing. In other words, police, judges, probation officers, and social workers have unintentionally come to associate higher criminal risk with certain racial groups—specifically people of color. New technologies such as risk assessments have been developed to intercept the individual discretion of practitioners. However, they too produce their share of unintended race-based consequences. How do we examine and address such problematic aspects of these technologies without completely abandoning our use of them?

The assessment of risk—an individual's likelihood of committing a new crime—can be an important aspect of pretrial release, sentencing, community supervision, and parole decisions. Some reformers are seeking to address over-incarceration and implicit bias through

the introduction of actuarial science ("big data") to the risk assessment process. As many as 60 data-driven risk assessment tools, diverse in form, length, and content, are currently in use across the United States. A key appeal of actuarial tools has been their ability, in the aggregate, to outperform professional judgment in terms of accuracy. In other words, data-driven predictions of criminal behavior are generally more accurate than subjective, professional predictions. Yet a lingering question remains: how fair are data-driven risk assessments when it comes to race?

There has been significant debate in the academic and popular press regarding the impact of actuarial risk assessments on racial disparities—specifically, whether risk assessments reduce racial disparities, exacerbate racial differences, or maintain the status quo. Some research has demonstrated that risk assessments could mitigate patterns of racial bias in decision-making. For example, a pattern of disproportionate pretrial detention among African-American juvenile defendants in one Colorado jurisdiction was eliminated following the introduction of a risk tool (Eaglin & Solomon, 2016).

Risk assessment proponents argue that data-driven risk assessment tools not only improve the accuracy of decisions but also can serve to effectively mitigate racial disproportionalities arising from implicit biases in laws, police practices, or the discretionary patterns of individual

R/NA

RISK / NEEDS ASSESSMENT
A PERSPECTIVES SPOTLIGHT

The problem goes beyond the existence of false positives—which are statistically inevitable with all algorithms—to specifically involve the type of data points used to define high- and low-risk categories

decision-makers. In particular, existing research suggests that disproportionalities in arrest and incarceration are especially prevalent in reference to low-level offenses—those cases where police and prosecutors exercise the most discretion—and the result is that many local jails spend millions to detain low-risk, low-level offenders (Golub, Johnson, & Dunlap, 2007; Natapoff, 2015). Risk assessment has the potential to make these inequities more transparent and to provide a compelling justification for limiting discretion (Rempel, Kerodal, Spadafore, & Mai, 2017).

Conversely, an argument can be made that risk tools may perpetuate racial disparities due to correlations between common risk factors and race. For example, unemployment, lack of education, and criminal history are said to have become proxies or stand-ins for race (Starr, 2015a; Harcourt, 2015; Horwitz, 2014). A report produced by a working group under the Obama Administration touches on this issue in discussing the creation of algorithms and use of machine learning for processing data. It notes that the knowledge, motives, and biases of an algorithm's author will affect the outcome of the recommendation engine. "[The] final computer-generated product or decision—used for everything from predicting behavior to denying opportunity—can mask prejudices while maintaining a patina of scientific objectivity." (Executive Office of the President, 2014, p. 46).

The complicated and often paradoxical relationship between risk assessment and racial disparities partly arises from differences in the total number of defendants of color and white defendants who are drawn into the justice system and subsequently assessed for risk in a given jurisdiction (also referred to as differential "base rates"). Even the best predictive algorithms will invariably misclassify some percentage of low-risk defendants as high risk (false positives) and some percentage of high-risk defendants as low risk (false negatives).

R/NA

RISK / NEEDS ASSESSMENT

A PERSPECTIVES SPOTLIGHT

Moreover, in the United States, where a disproportionate number of people of color get arrested and prosecuted compared to their white counterparts, the use of any single algorithm to predict risk will consequently yield more false positives for defendants of color. In other words, if more defendants of color get assessed than whites, then a greater number of those defendants will be misclassified as high risk. This problem was recently demonstrated in one jurisdiction in Florida, where black defendants were disproportionately classified as high risk but were not actually re-arrested (Angwin, J, Larson, J, Mattu, S, & Kirchner, L., 2016).

The problem goes beyond the existence of false positives—which are statistically inevitable with all algorithms—to specifically involve the type of data points used to define high- and low-risk categories in criminal justice. Most risk-and-need factors that algorithms associate with recidivism are not racially neutral. This means that defendants of color may fare worse in terms of the individual results of a risk assessment tool due to historical racial bias, socioeconomic conditions, and crime policy trends that are beyond any individual defendant's control (Spohn, 2015). For instance, hot spots policing results in higher arrest rates and longer average criminal records for black defendants, with criminal background being the single

strongest factor in most risk assessment tools (Harcourt, 2015). Needs that are associated with recidivism risk, such as homelessness and unemployment, are also disproportionately found in non-white populations. The bottom line is that although most risk assessment tools are designed to be “color blind,” involvement in the justice system is not.

Ultimately, recognizing and developing fair risk assessment tools is not merely a mathematical exercise. Instead, it will be a question of values. One value is maximized racial equity and another public safety, which raises the possibility of a trade-off between “fairness” and perceived public safety (Corbett-Davies, Pierson, Feller, Goel, & Huq, 2017).

Before we can assess such a trade-off, we need to define fairness. Does a fair tool simply mean that the rate of re-arrest is equivalent within categories despite race (i.e., all high-risk individuals have a 60% likelihood of re-arrest)? Or does it mean the tool is equally statistically accurate (i.e., doesn't have more false positives or false negatives for each racial group)? Or, finally, does fairness mean that the percentage of people deemed high or low risk is the same across all racial groups, resulting in greater cross-group parity?

R/NA

RISK / NEEDS ASSESSMENT

A PERSPECTIVES SPOTLIGHT

Balancing these values in the exercise of risk assessment should be pursued with the utmost transparency and with eyes toward both the challenges of real-world implementation and the demonstrably harmful effects of mass incarceration. To this end, we propose several strategies for consideration.

POLICY STRATEGIES

The use of actuarial risk tools to reduce unnecessary interventions (e.g., booking, detention, imprisonment) may increase the equity of the system. Application of risk tools specifically to pretrial release in lieu of money bond setting is one example. This strategy was implemented statewide in New Jersey in 2016 and is under consideration by several other jurisdictions (Rabner, 2017). Although not widely acknowledged in the policy sphere, risk assessment may also support more equitable approaches to policing. Brief, records-based assessments may be practical tools for officers in some jurisdictions and can serve to counter existing police assumptions or biases about who is high risk and, therefore, should be subject to custodial arrest rather than being diverted or issued a citation (Picard-Fritsche, Spadafore, Lebron, & Jensen, 2017).

Use of comprehensive risk-need assessment tools to build new (or improve existing) alternative-to-incarceration

programs has strong potential for reducing justice system involvement by offering better front-end diversion alternatives and reducing recidivism for re-entry populations. For instance, if homelessness is a strong predictor for re-arrest, as it is in many urban jurisdictions, then programs that provide support in finding stable housing will be more effective. Thirty years of risk-need research provides specific guidance on what works in recidivism reduction.

TECHNICAL STRATEGIES

Technocratic solutions—i.e., revisions to the content or structure of risk assessment tools themselves—may also help maximize equity in risk assessment. Currently, most risk assessment tools are focused on calculating the effect of *individual* risk factors (e.g., age, criminal record, substance use patterns) on recidivism, but this focus is not a foregone conclusion. The neighborhood in which an individual is arrested, the jurisdiction in which a trial takes place, and/or the judge who oversees the case may also affect risk profile. Such strategies may be used by tool developers to increase equity if, for example, being arrested in a neighborhood subject to proactive policing is calculated as a factor that should mitigate as opposed to inflate risk.

Designing risk assessment tools for a specific context or group is another

R/NA

RISK / NEEDS ASSESSMENT

A PERSPECTIVES SPOTLIGHT

technocratic approach to increasing equity. Local validation, or specification of a risk algorithm at the jurisdiction level, has long been promoted as best practice in the field (Miller & Lin, 2007). Design of group-specific algorithms that account for historical bias against specific racial groups (e.g., an algorithm exclusively for black defendants) could also lessen bias compared with “race neutral” tools that define risk the same for all racial groups. Similar strategies have been employed to account for gender disparities in risk prediction. These strategies are more complex and controversial when applied to race, although it is notable that “culturally competent” risk assessment tools have been proposed in some international contexts (Hannah-Moffat & Maurutto, 2010).

In conclusion, creating a context for the ethical and effective use of risk assessment in criminal justice may require us to question some of our fundamental assumptions about both fairness and science. First, we must wrestle with the notion that race neutrality may not amount to fairness, given our history. Instead, risk assessments should be considered fair and unbiased to the extent that they are responsive to a reality in which criminal risk is shaped by race group membership. Second, we must admit that risk assessment tools, like other strategies drawing on science, do not in

and of themselves guarantee fairness and an ethical application of policy. Reform-minded stakeholders in jurisdictions across the country should begin deliberately, and collaboratively, considering the trade-offs built into their risk assessment tools prior to implementation. Risk assessments may address mass incarceration as a social problem only to the degree that they are explicitly employed toward this goal by researchers and policymakers alike.

REFERENCES

- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May). Machine bias: There’s software used across the country to predict future criminals. And it’s biased against blacks. *ProPublica*. Retrieved from <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Austin, J. (2003). The proper and improper use of risk assessment in corrections. *Federal Sentencing Reporter*, 16(3), 193-1998. Retrieved from www.jfa-associates.com/publications/pcras/proper%20userand%20misuse%20of%20risk.pdf
- Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, Z. (2017, February). Algorithmic decision making and the cost of fairness. Retrieved from <https://arxiv.org/pdf/1701.08230.pdf>.
- Executive Office of the President. (2014). *Big data: Seizing opportunities, preserving values*. Washington, DC: Author. Retrieved from: https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf
- Hannah-Moffat, K., & Maurutto, P. (2010). Restructuring pre-sentence reports: Race, risk, and the PSR. *Punishment and Society*. 12(2), 262-286. Retrieved from <http://pun.sagepub.com/content/12/3/262.abstract>.

R/NA

RISK / NEEDS ASSESSMENT

A PERSPECTIVES SPOTLIGHT

Harcourt, B. E. (2015). Risk as a Proxy for Race. *Federal Sentencing Reporter*, 27(4): 237-243.

Horwitz, S. (2014, August 2). Eric Holder: Basing sentences on data analysis could prove unfair to minorities. *The Washington Post*. Retrieved from: https://www.washingtonpost.com/world/national-security/us-attorney-general-eric-holder-urges-against-data-analysis-in-criminal-sentencing/2014/08/01/92d0f7ba-1990-11e4-85b6-c1451e622637_story.html.

Miller, J., & Lin, J. (2007). Applying a generic juvenile risk assessment instrument to a local context: Some practical and theoretical lessons. *Crime & Delinquency*, 53(4), 552-580.

Natapoff, Alexandra. (2015). Misdemeanors. *Annual Review of Law and Social Science*, (11), 255-267. Available at SSRN: <https://ssrn.com/abstract=3689015> or <http://dx.doi.org/10.1145/annurev-lawsocsci-120814-121742>.

Rabner, S. (2017, January 9). Chief justice: Bail reform puts N.J. at the forefront of fairness | Opinion. *New Jersey Star Ledger*. Available at: <http://www.njcourts.gov/criminal/cjr/CJ%20Sunday%20Star-Ledger%20Column%20-%20Bail%20Reform%20Puts%20NJ%20at%20the%20Forefront%20of%20Fairness%20-%2001-08-17%20-%20posted%2001-09-17.pdf>.

Rempel, M, Kerodal, A, Spadafore, J., & Mai, C. (2017). Jail in New York City: Evidence-based opportunities for reform. New York, NY: Center for Court Innovation.

Spohn, C. (2015). Race, crime, and punishment in the twentieth and twenty-first centuries. *Crime and Justice*, 44(1), 49-97.

Starr, S.B. (2015). The risk assessment era: An overdue debate. *Federal Sentencing Reporter*, 27, 205-206.

ABOUT THE AUTHORS

SARAH PICARD-FRITSCH is the deputy director of Research-Practice Strategies at the Center for Court Innovation. She currently co-leads the Center's research and technical assistance effort to reduce the use of jail incarceration nationally, funded by the MacArthur Foundation. In this capacity, she has been integral to the Center's work examining the drivers of racial disparities in incarceration. She is also the principal investigator on several studies examining the use of actuarial risk assessment in justice-system settings. She can be reached at sfritsch@nycourts.gov.

JULIAN ADLER is the director of Research-Practice Strategies at the Center for Court Innovation, spanning the agency's three primary areas of work: social science research, local operating programs (New York and New Jersey), and national technical assistance. Julian directs the Center's work on the John D. and Catherine T. MacArthur Foundation's Safety and Justice Challenge. He also leads the Center's work on the Price of Justice. He is a New York State licensed clinical social worker (LCSW) and attorney. He can be reached at jadler@nycourts.gov.

YOLAINE MENYARD is the associate director of Research-Practice Strategies at the Center for Court Innovation, which synergizes the agency's three primary areas of work: research, operating programs, and expert assistance. Ms. Menyard works on the development of holistic evidence-informed and evidence-generating practices, including assessment instruments and short-term interventions. Ms. Menyard is a licensed master social worker in the State of New York. She completed her undergraduate studies at the University of Notre Dame, and received her MSW from the University of Chicago. She can be reached at ymenyard@nycourts.gov.

SHUBHA BALA is the product manager for Technology and Innovation at the Center for Court Innovation. In this capacity, she is responsible for identifying, implementing, and evaluating technology that helps Center programs and their clients. She holds a bachelor of science in electrical engineering from the University of Toronto and a masters in international affairs from Columbia University's School of International and Public Affairs. She can be reached at sbala@nycourts.gov.